

LC-HRMS-Database Screening Metrics for Rapid Prioritisation of Samples to Accelerate the Discovery of Structurally New Natural Products

Jioji N. Tabudravu,^{*,†,‡} Léonie Pellissier,[‡] Alan James Smith,[‡] Karolina Subko,[‡] Caroline Autréau,[‡] Klaus Feussner,[§] David Hardy,[⊥] Daniel Butler,^{||} Richard Kidd,[∇] Edward J. Milton,[○] Hai Deng,[‡] Rainer Ebel,[‡] Marika Salonna,[∠] Carmela Gissi,^{∠Ω} Federica Montesanto,[#] Sharon M. Kelly,[□] Bruce F. Milne,[◇] Gabriela Cimpan,[○] Marcel Jaspars^{*,‡}

[†] School of Forensic and Applied Sciences, Faculty of Science & Technology, University of Central Lancashire, Preston, Lancashire, PR1 2HE, UK.

[‡] Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, AB24 3UE, Scotland, UK.

[§] Institute of Applied Sciences, Faculty of Science, Technology and Environment, University of the South Pacific, Laucala Campus, Private Mail Bag, Suva, Fiji Islands.

[⊥] Thermo Fisher Scientific, Altrincham Business Park, 1 St George's Court, Altrincham WA14 5TP, UK.

^{||} 31 Hobbs Close, OX14 3UX, Oxford, UK.

[∇] Publisher, Data & Databases, Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, CB4 0WF, UK.

[○] Advanced Chemistry Development, UK Ltd. Venture House, Arlington Square, Downshire Way, Bracknell, Berks. RG12 1WA, UK.

[◇] CFisUC, Department of Physics, University of Coimbra, Rua Larga, 3004-516, Coimbra, Portugal

[∠] Department of Biosciences, Biotechnologies and Biopharmaceutics, University of Bari "A. Moro", Via Orabona 4, 70125 Bari, Italy.

^{∠Ω} IBIOM, Istituto di Biomembrane, Bioenergetica e Biotechnologie Molecolari, CNR, Via Amendola 165/A, 70126 Bari, Italy.

[#] Department of Biology - LRU CoNISMa, University of Bari, Via Orabona 4, 70125, Bari, Italy

[□] Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G128QQ, UK.

Table of Contents

Figure/Table	Page
Part A. Spectroscopic and other data	4
Figure S1. ¹ H NMR spectrum of tavarua deoxyriboside A (4)	4
Figure S2. Edited HSQC NMR spectrum of tavarua deoxyriboside A (4)	4
Figure S3. COSY NMR spectrum of tavarua deoxyriboside A (4)	5
Figure S4. HMBC NMR spectrum of tavarua deoxyriboside A (4)	5
Figure S5. HRESIMS (ES ⁺) spectrum of tavarua deoxyriboside A (4)	6
Figure S6. MS/MS data of tavarua deoxyriboside A (4)	6
Figure S7. 2D ROE NMR spectrum for tavarua deoxyriboside A (4)	7
Figure S8. Minimised conformation structure for tavarua deoxyriboside A (4)	8
Figure S9. HRESIMS of tavarua deoxyriboside B (5) plus major fragment	9
Figure S10. ¹ H NMR spectrum tavarua deoxyriboside B (5)	9
Figure S11. ¹ H NMR spectrum of eudistomin Z1 (1)	10
Figure S12. COSY NMR spectrum of eudistomin Z1 (1)	10
Figure S13. HSQC NMR spectrum of eudistomin Z1 (1)	11
Figure S14. HMBC NMR spectrum of eudistomin Z1 (1)	11
Figure S15. HRESIMS (ES ⁺) of eudistomin Z1 (1)	12
Figure S16. Predicted against experimental ¹³ C NMR data for eudistomin Z1 (1)	12
Figure S17. ¹ H NMR spectrum of eudistomin Z2 (2)	13
Figure S18. COSY NMR spectrum of eudistomin Z2 (2)	13
Figure S19. HSQC NMR spectrum of eudistomin Z2 (2)	14
Figure S20. HMBC NMR spectrum of eudistomin Z2 (2)	14
Figure S21. HRESIMS (ES ⁺) of eudistomin Z2 (2)	15
Figure S22. Predicted against experimental ¹³ C NMR data for eudistomin Z2 (2)	15
Figure S23. Icons for calculating metrics data	16
Table S24. Indexed masses (total) after applying scripted IX	16
Table S25. Selection of total ions after applying IX – deconvoluted	17
Table S26. Non identified masses after screening of PharmaSeaDB and MatrixDB	18
Table S27. Metric data arranged in decreasing order of sample novelty	19
Figure S28. PharmaSeaDB	20
Table S29. Metric data for Eden-6-SPE-50 in PharmaSeaDB	20
Table S30. Masses identified/not identified in PharmaSeaDB for eden-6-SPE-50	21-23
Table S31. Masses not identified in PharmaSeaDB for eden-6-SPE-50	23-25
Table S32. Masses not identified in PharmaSeaDB for tavarua-2-SPE-50	25
Table S33. Masses identified/not identified in PharmaSeaDB for tavarua-2-SPE-50	25-26
Figure S34. Metric data for Tavarua-2-SPE-50 in PharmaSeaDB	27
Figure S35. Parameters used by IntelliXtract (IX.mcr) for processing of LC-HRMS data	27-32
Table S36. Complexity metric data arranged from highest to lowest	33
Table S37. Diversity metric data arranged from highest to lowest	34
Table S38. Fijian marine sponges and tunicates used in this study	35
Figure S39. ¹ H NMR for tavarua-2-SPE-50	35
Figure S40. ¹ H NMR for tavarua-8-SPE-100	36
Figure S41. Metric data for tavarua-8-100	36
Table S42. Unidentified masses for tavarua-8-100	37-38
Figure S43. Metric data for tavarua-2-50, mass range of 0.2 ppm	38
Table S44. New masses in tavarua-2-50, mass range of 0.2 ppm	39
Figure S45. Total ion chromatogram of tavarua-2-50 peak at 8.59 min (<i>m/z</i> =522)	40
Figure S46. UV spectrum for compound 4	40
Figure S47. UV spectrum for compound 5	41
Figure S48. IR spectrum for compound 4	41
Figure S49. ECD spectrum for compound 4	42
Figure S50. UV spectrum for compound 1	42

Figure S51. UV spectrum for compound 2	43
Part B. Computational calculation of ECD	43
Figure S52-58. ECD overlay of cal. vs exp. for tavarua deoxyribose A (stereoisomers)	44-46
Figure S59. Structures of 8 possible stereoisomers of 4	46
Part C – Ascidian taxonomy	47-56
Figure S60. Tava8 morphology	51
Figure S61. Tava2 morphology	51
Figure S62. 18S rDNA phylogenetic tree of Aplousobranchia	52
Figure S63. Phylogenetic tree of Aplousobranchia – mitochondrial COI gene	53
Table S64. List of COI gene sequences included in the phylogenetic tree	54
Table S65. List of the 18S rDNA sequences	55
Table S66. Abbreviation of the ascidian genera analysed.	56

Part A

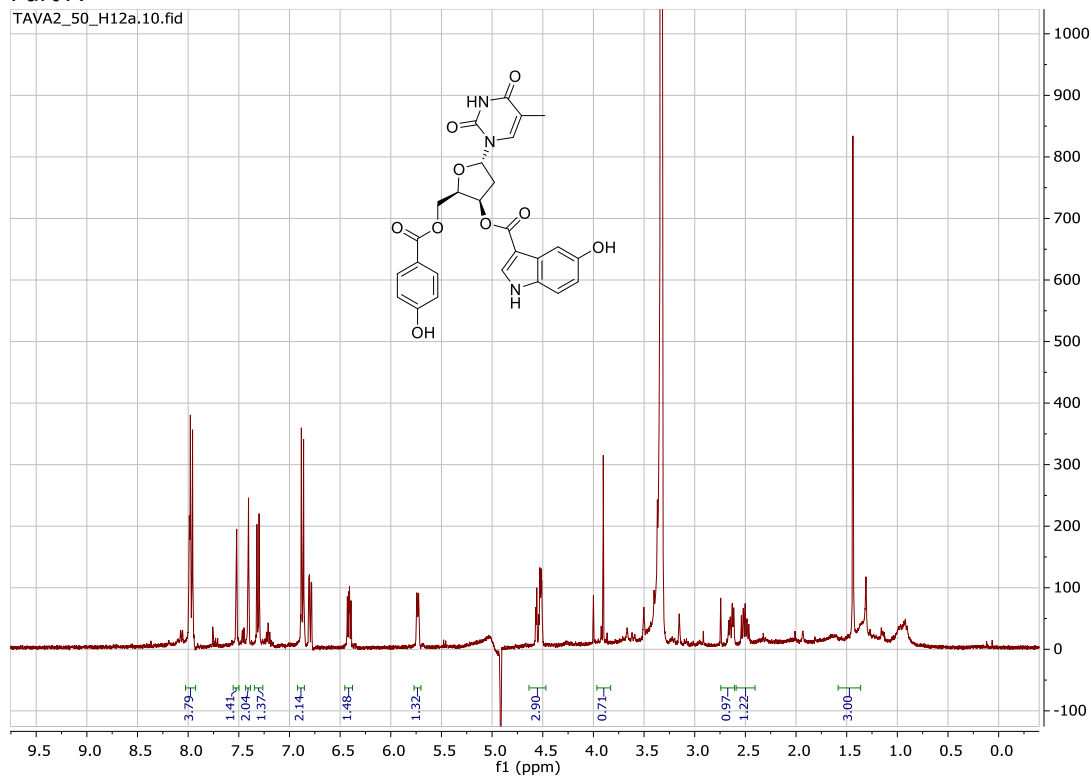


Figure S1. ^1H NMR spectrum (600 MHz, CD_3OD) for tavarua deoxyriboside A (**4**)

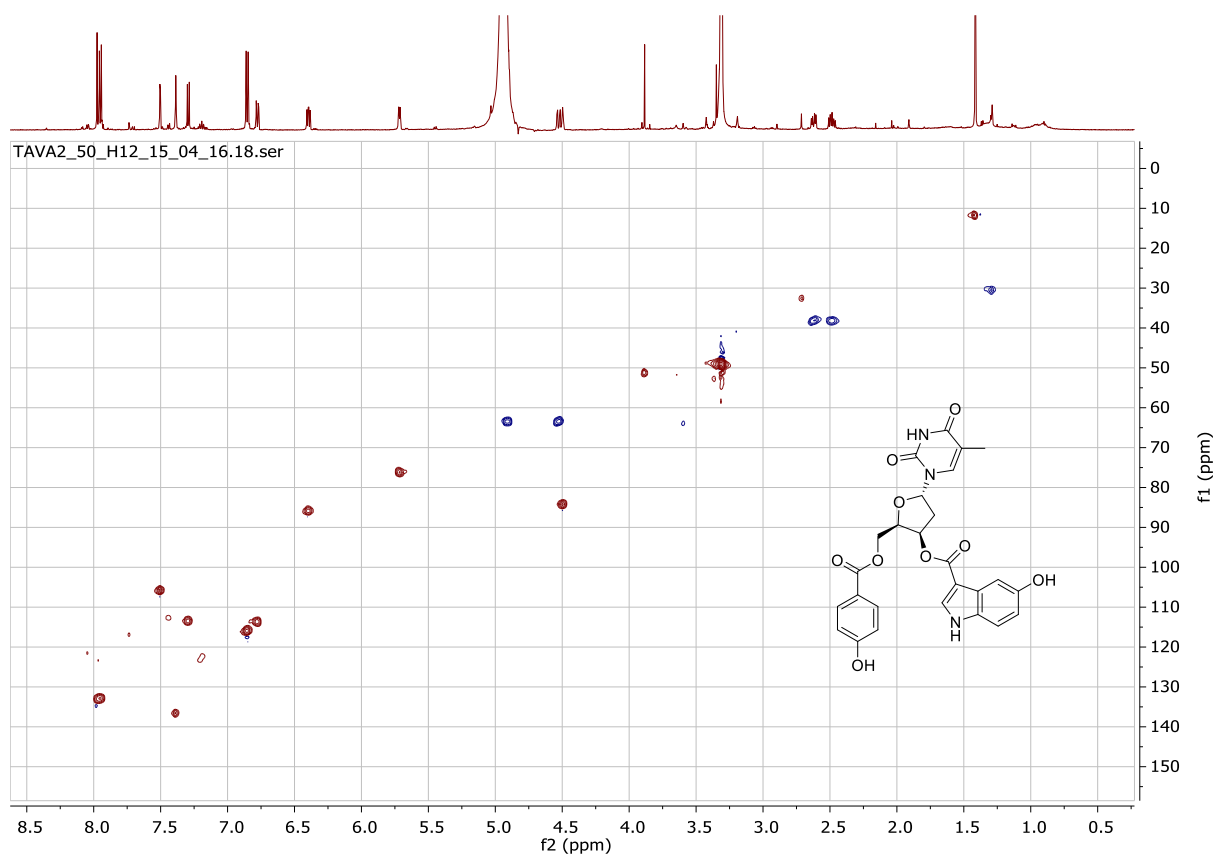


Figure S2. Edited HSQC NMR (600 MHz, CD_3OD) spectrum of tavarua deoxyriboside A (**4**)

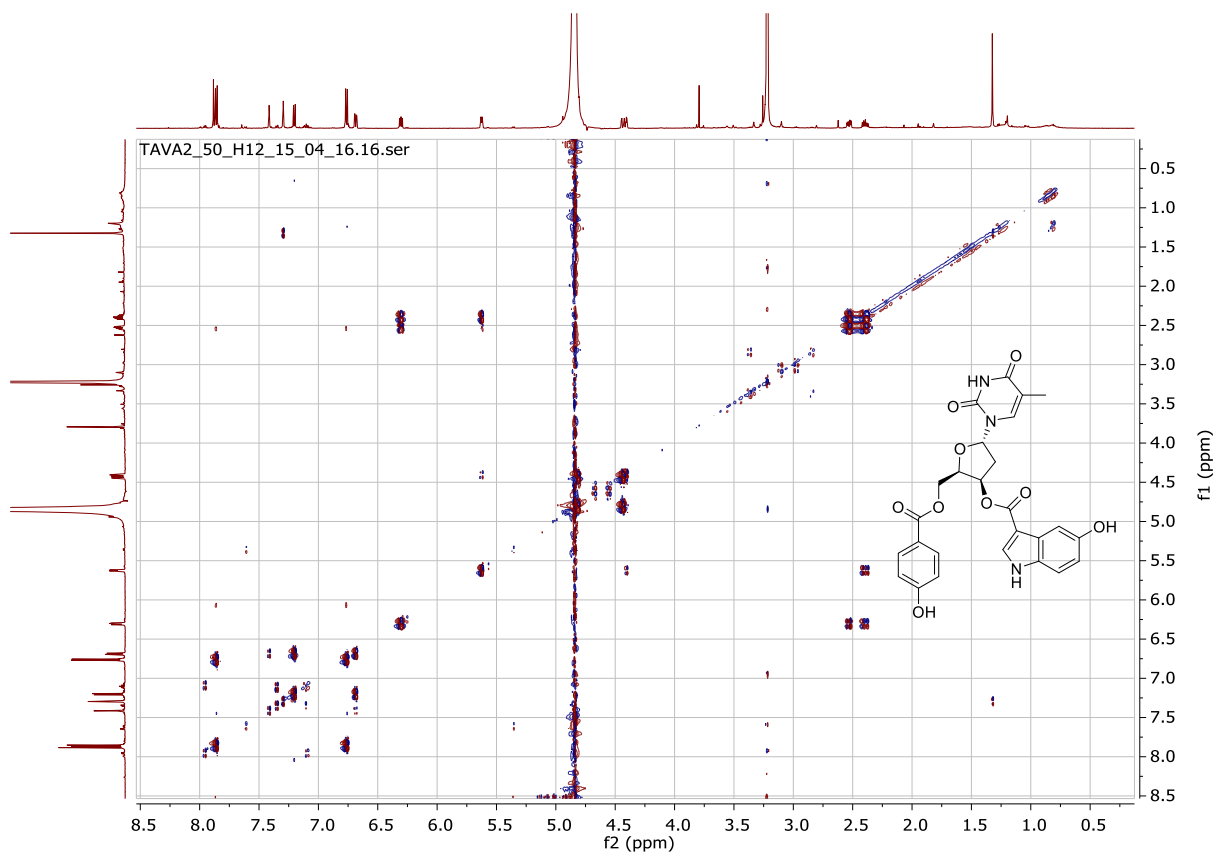


Figure S3. COSY NMR (600 MHz, CD₃OD) spectrum of tavarua deoxyriboside A (4)

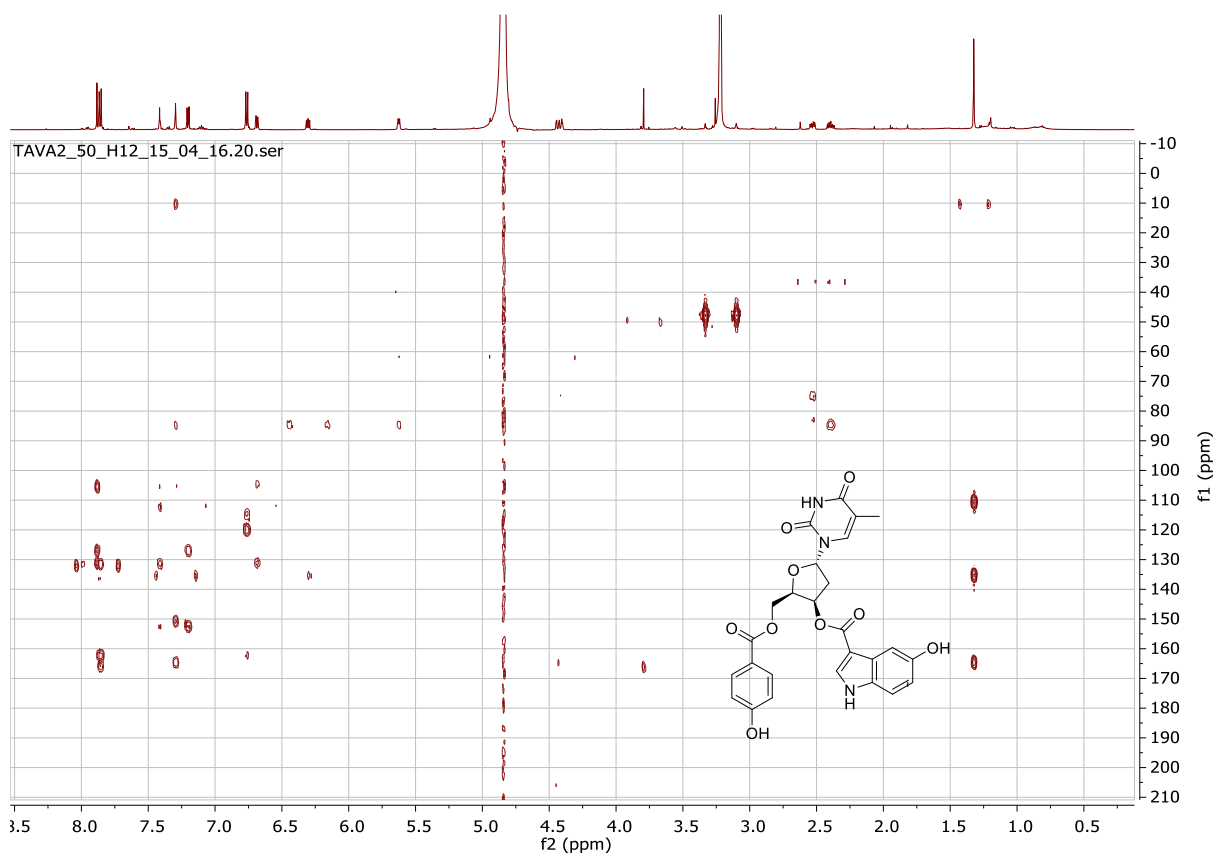


Figure S4. HMBC NMR (600 MHz, CD₃OD) spectrum of tavarua deoxyriboside A (4)

TAVA250H12 #391 RT: 8.81 AV: 1 NL: 3.73E7
F: FTMS + p ESI Full ms [100.00-2000.00]

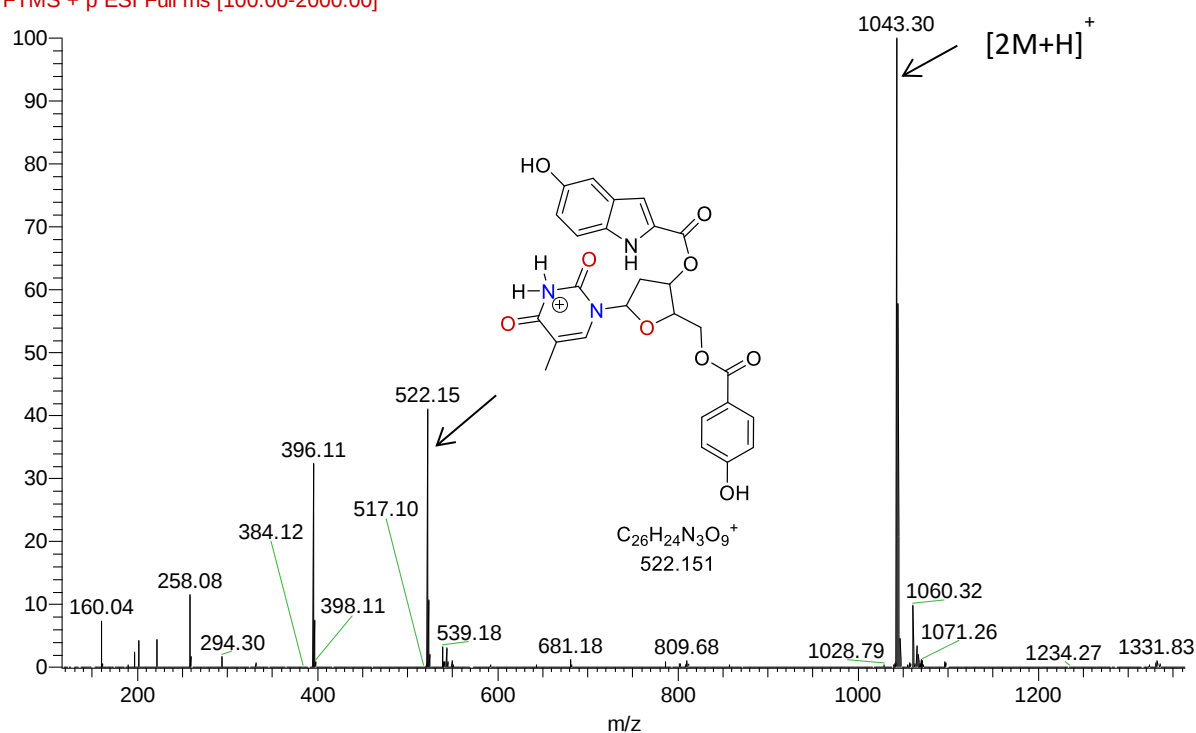


Figure S5. HRESIMS (ES⁺) spectrum of tavarua deoxyriboside A (4)

TAVA250H12 #397 RT: 8.94 AV: 1 NL: 6.12E6
F: FTMS + p ESI Full ms [100.00-2000.00]

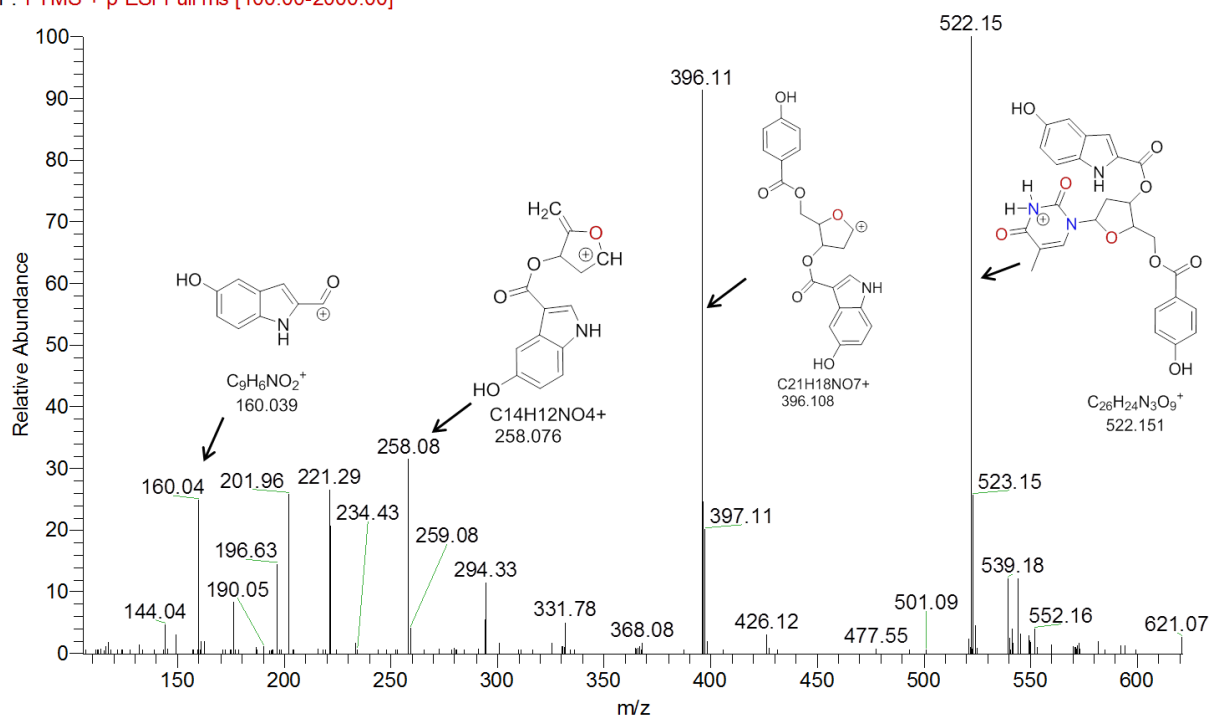


Figure S6. MS/MS data of tavarua deoxyriboside A (4).

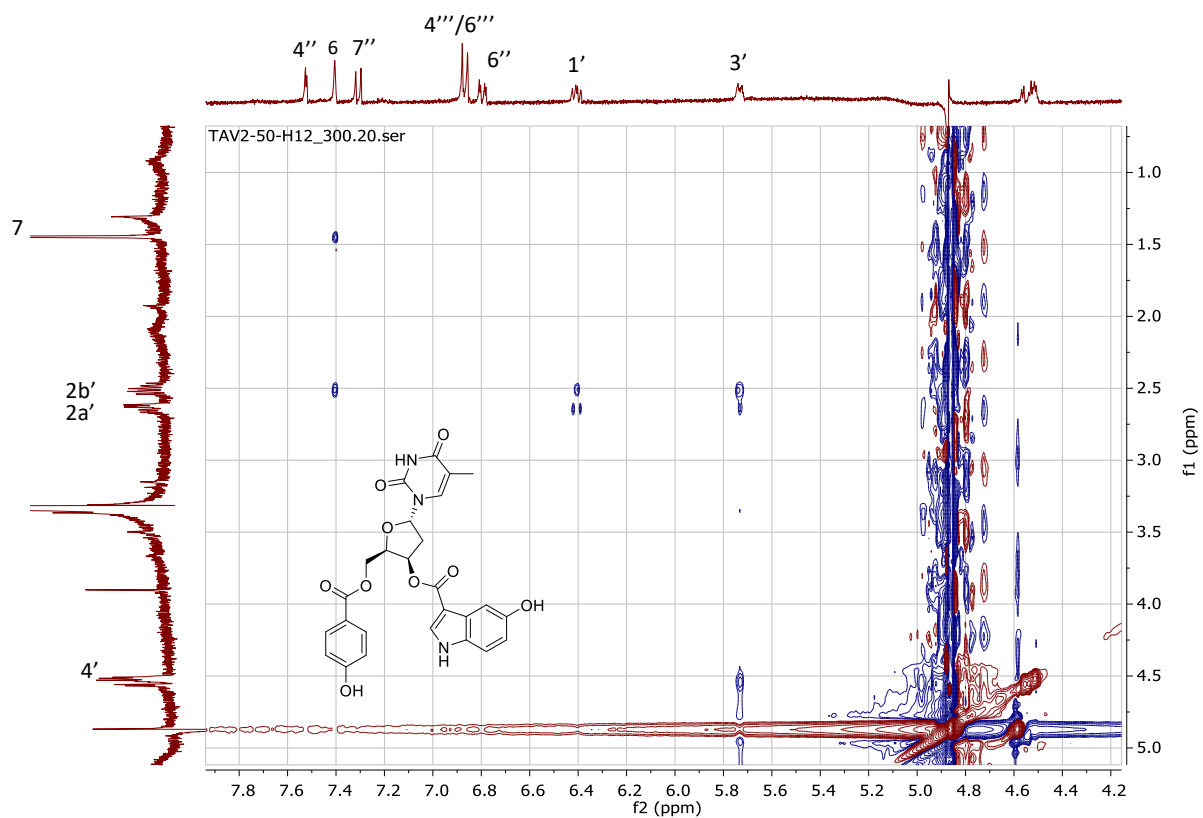


Figure S7. 2D ROESY NMR (600 MHz, 300 ms, CD₃OD) spectrum of tavarua deoxyribose A (**4**)

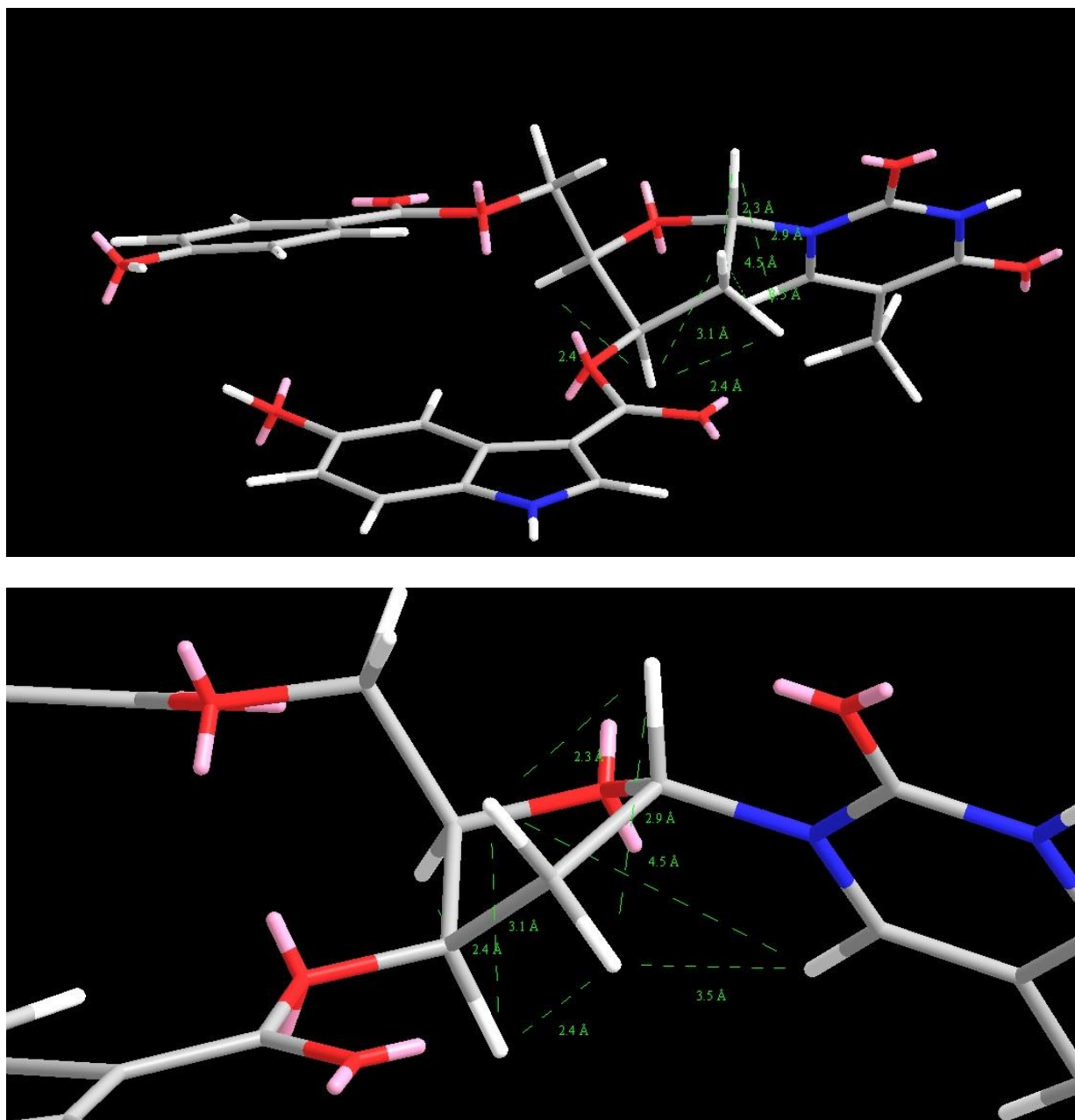


Figure S8. Minimized energy conformation structure for tavarua deoxyriboside A (**4**) showing observed ROE correlations suggesting an α bond between the deoxyribose unit (C-1') and the thymidine unit at N-1. Bottom figure shows expanded view for clarity (refer to Figure S7) for ROE correlations.

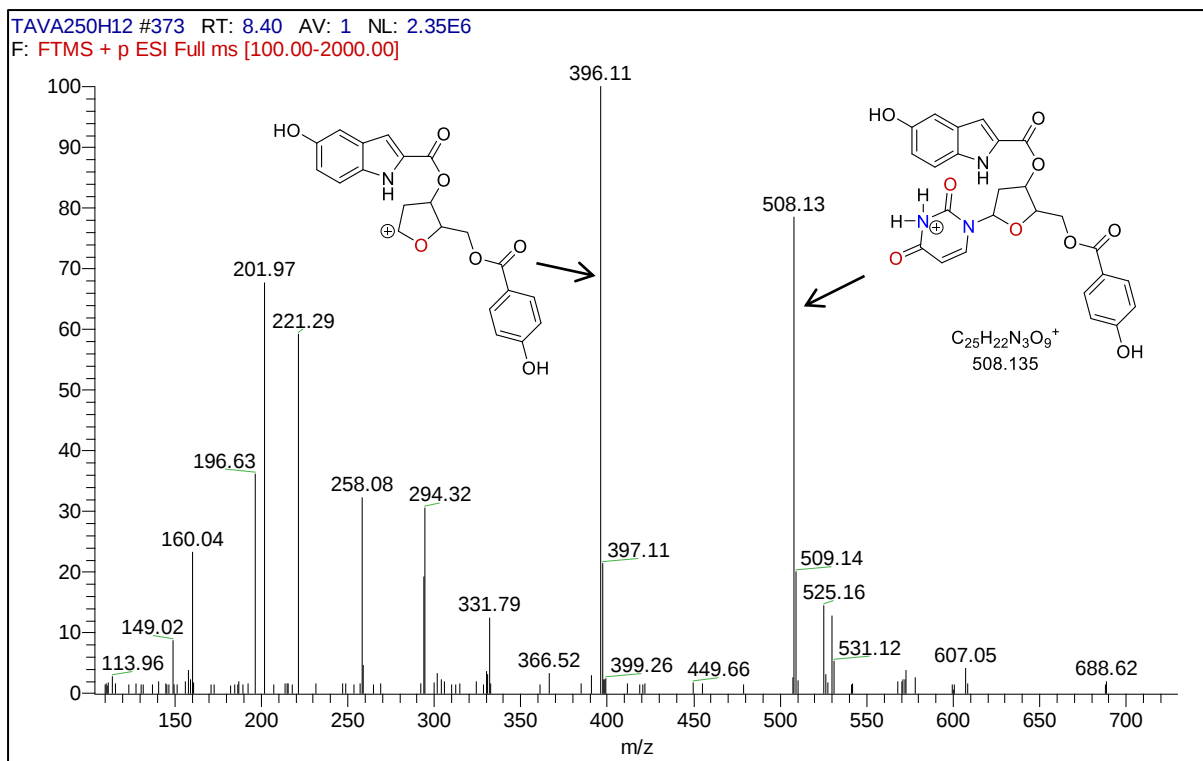


Figure S9. HRESIMS of tavarua deoxyriboside B (5) plus major fragment.

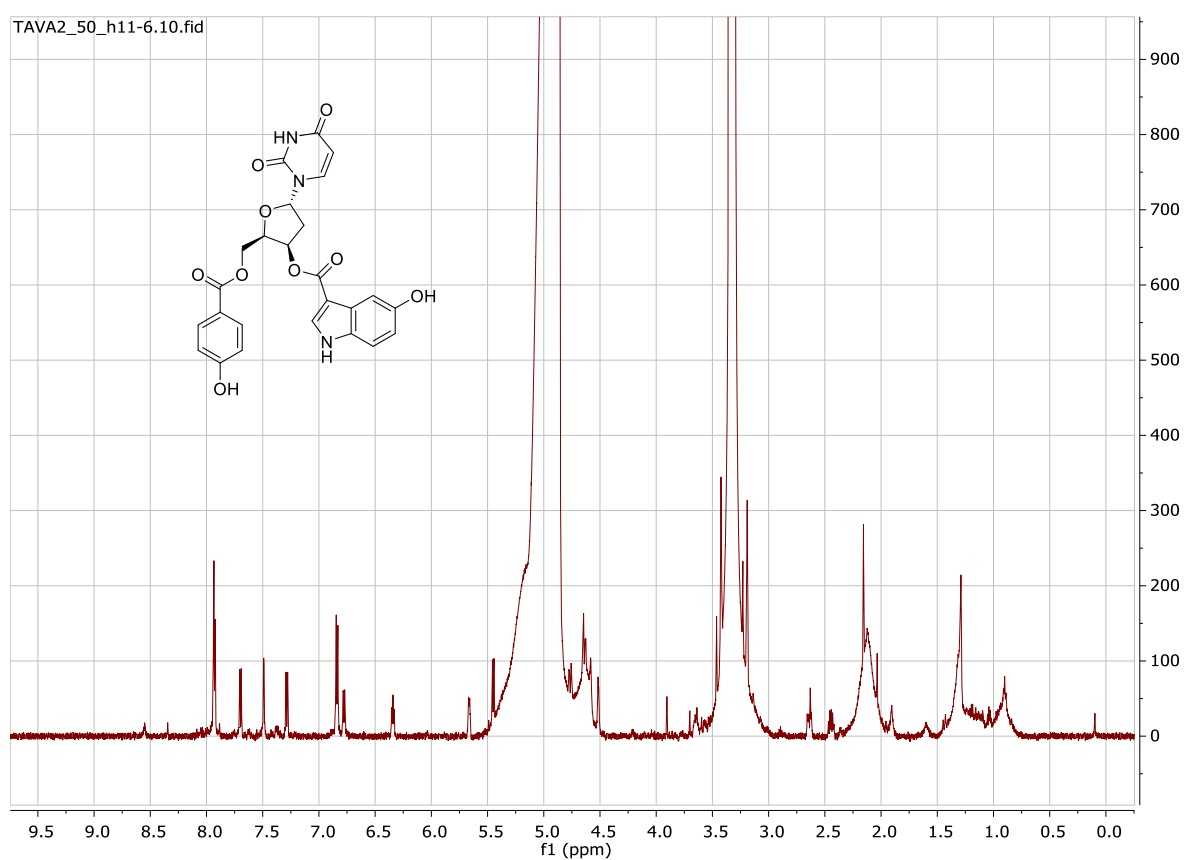


Figure S10. ^1H NMR (600 MHz, CD_3OD) spectrum of tavarua deoxyriboside B (5)

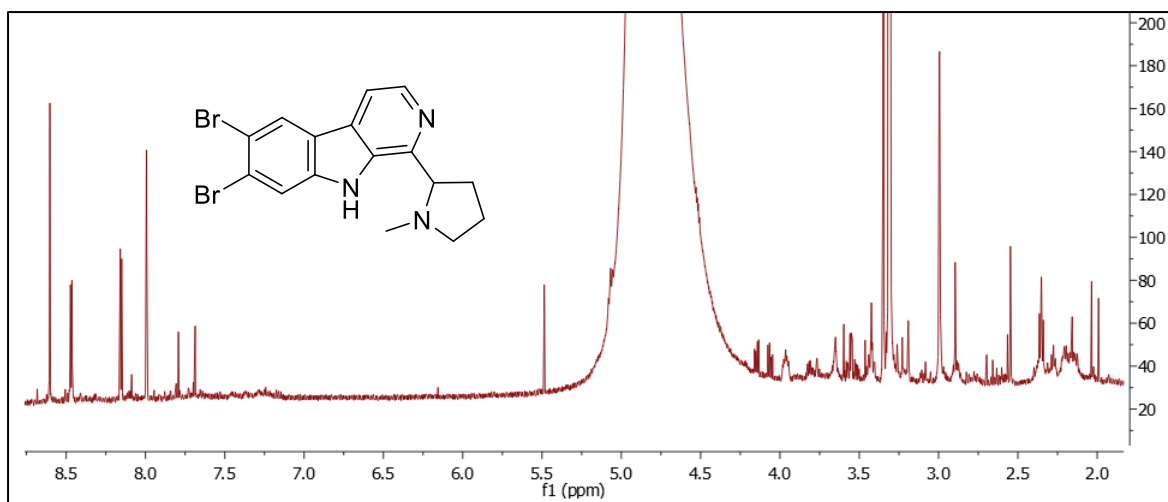


Figure S11. ^1H NMR (600 MHz, CD_3OD) spectrum of eudistomin Z1 (**1**)

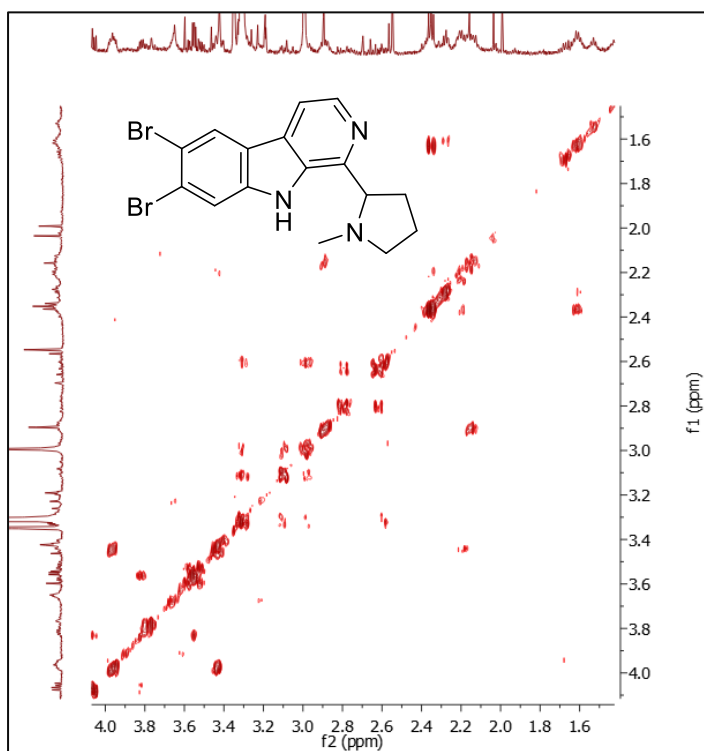


Figure S12. COSY NMR (600 MHz, CD_3OD) spectrum of eudistomin Z1 (**1**)

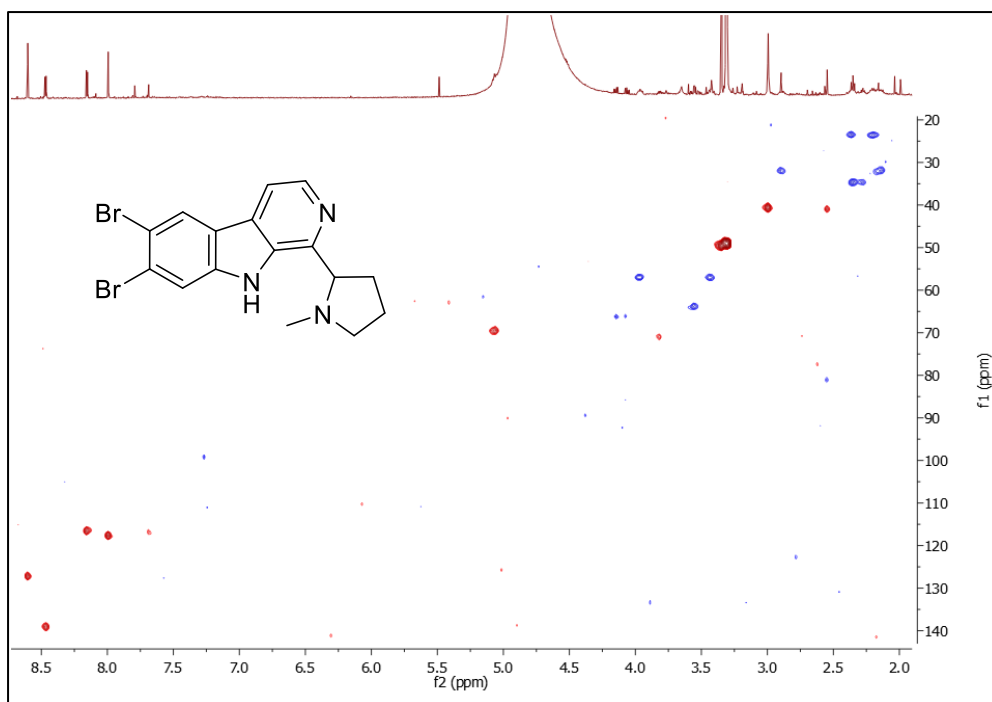


Figure S13. HSQC NMR (600 MHz, CD₃OD) spectrum of eudistomin Z1 (**1**)

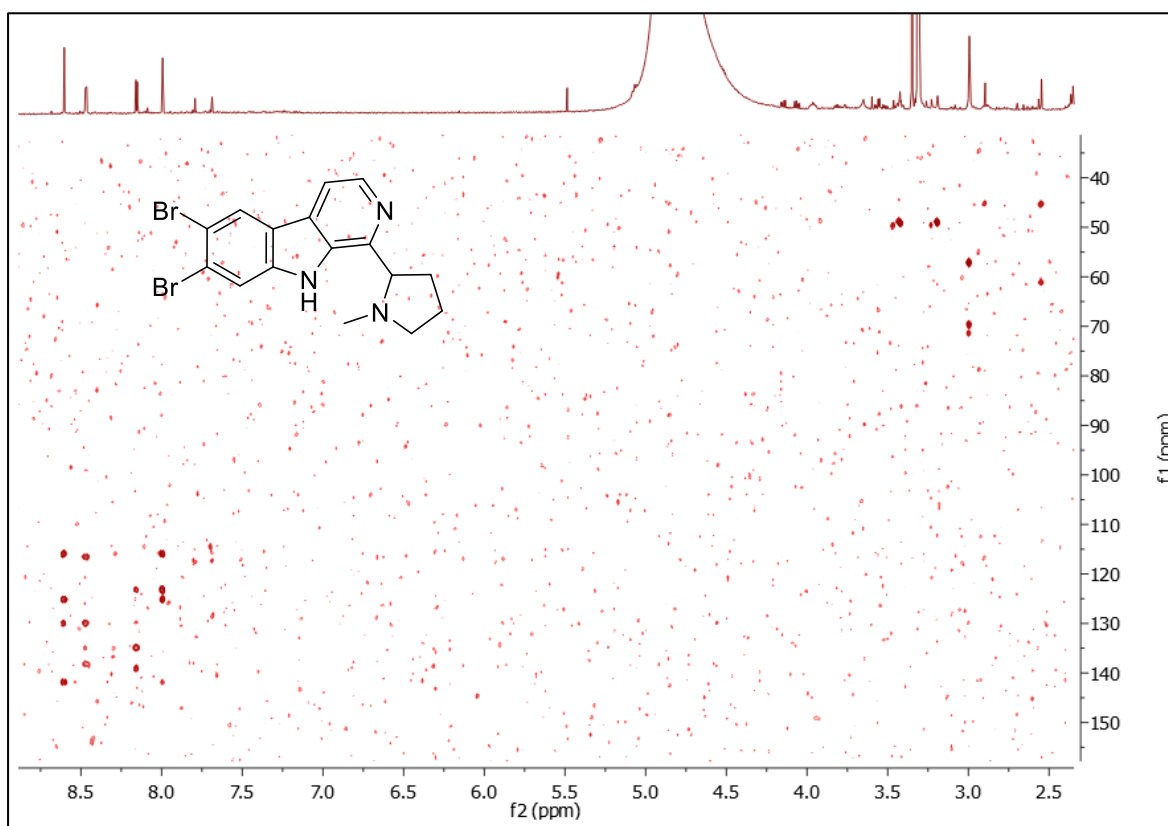


Figure S14. HMBC NMR (600 MHz, CD₃OD) spectrum of eudistomin Z1 (**1**)

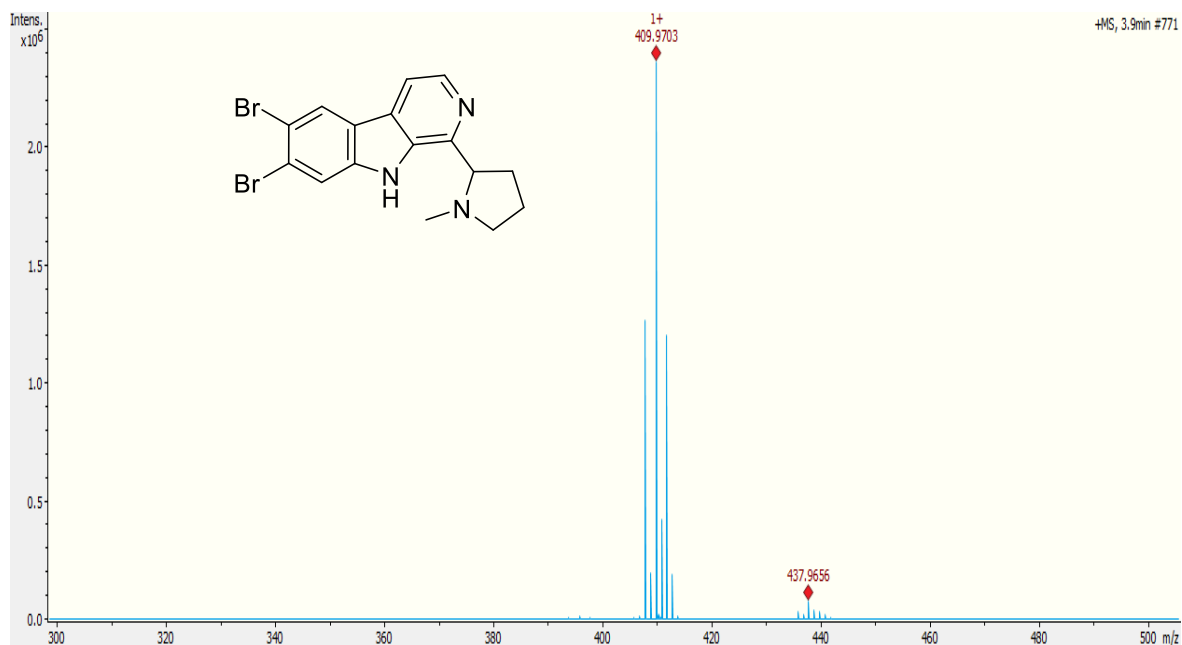


Figure S15. HRESIMS (ES⁺) of eudistomin Z1 (**1**)

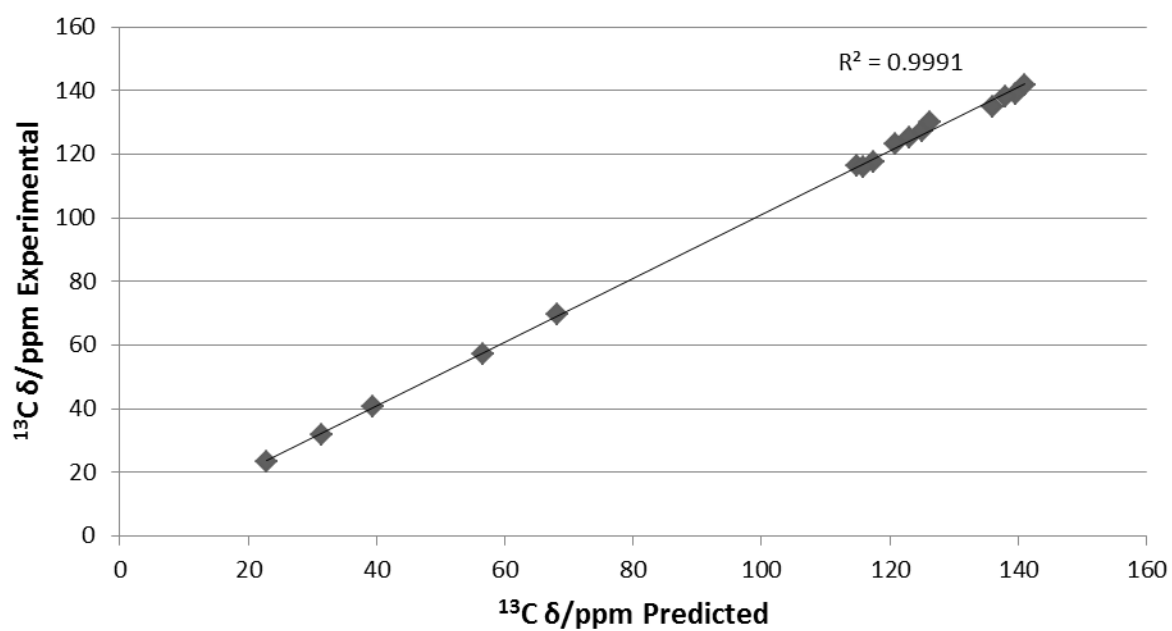


Figure S16. Predicted against experimental ¹³C NMR data for eudistomin Z1 (**1**)

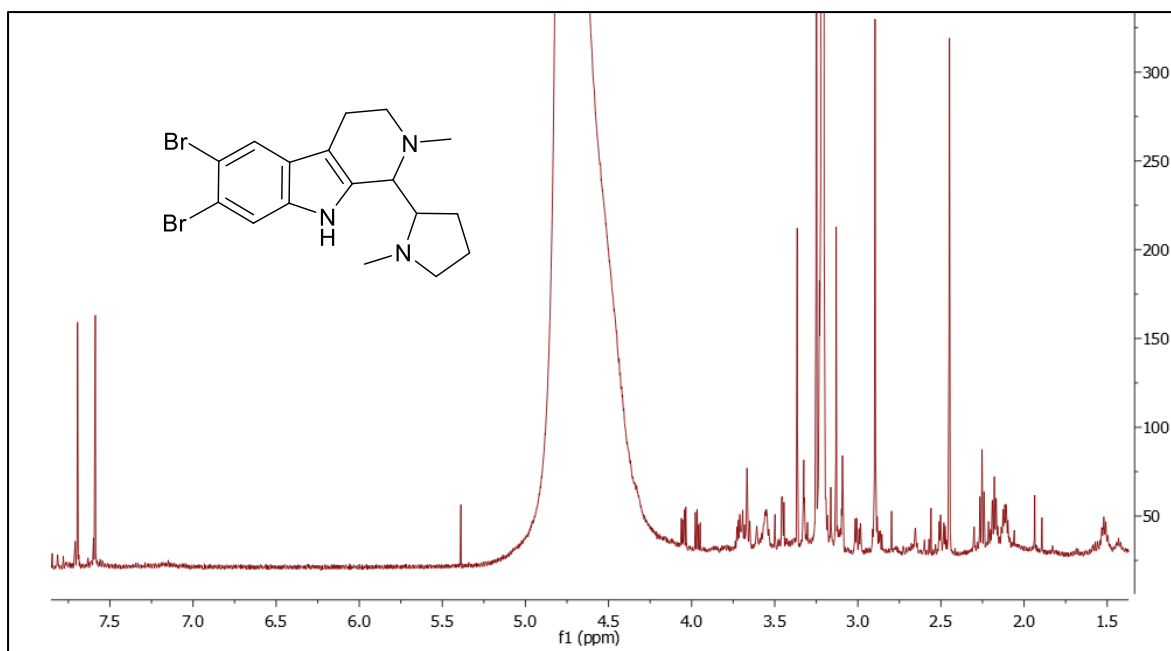


Figure S17. ^1H NMR (600 MHz, CD_3OD) spectrum of eudistomin Z2 (**2**)

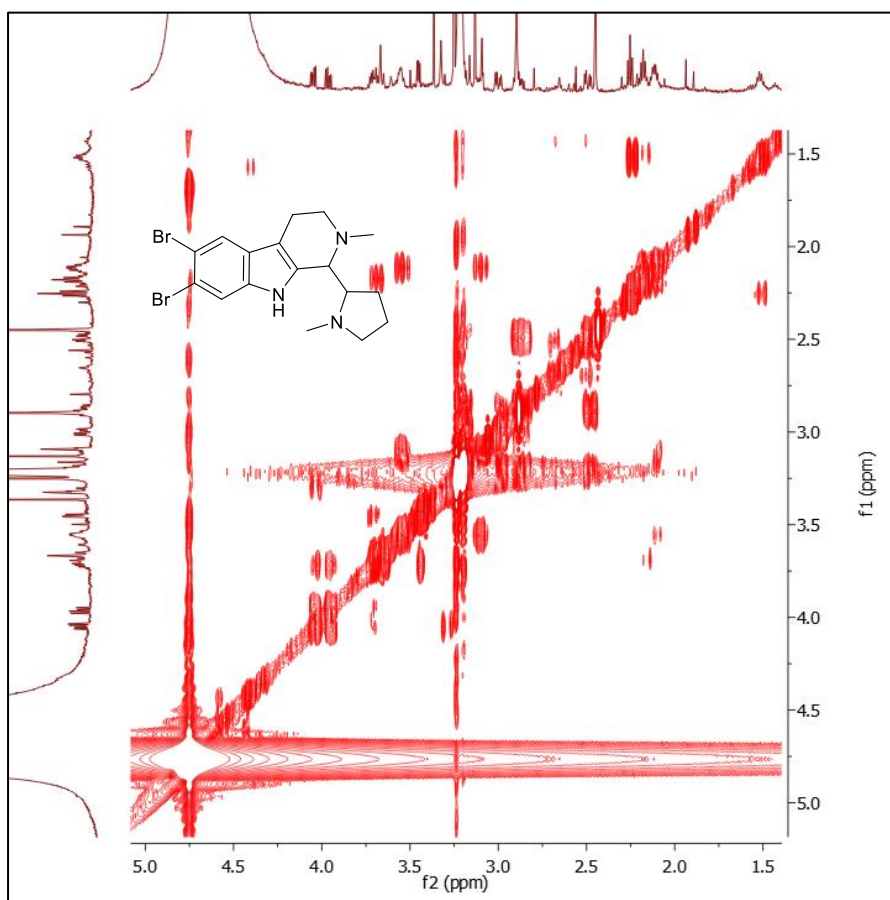


Figure S18. COSY NMR (600 MHz, CD_3OD) spectrum of eudistomin Z2 (**2**)

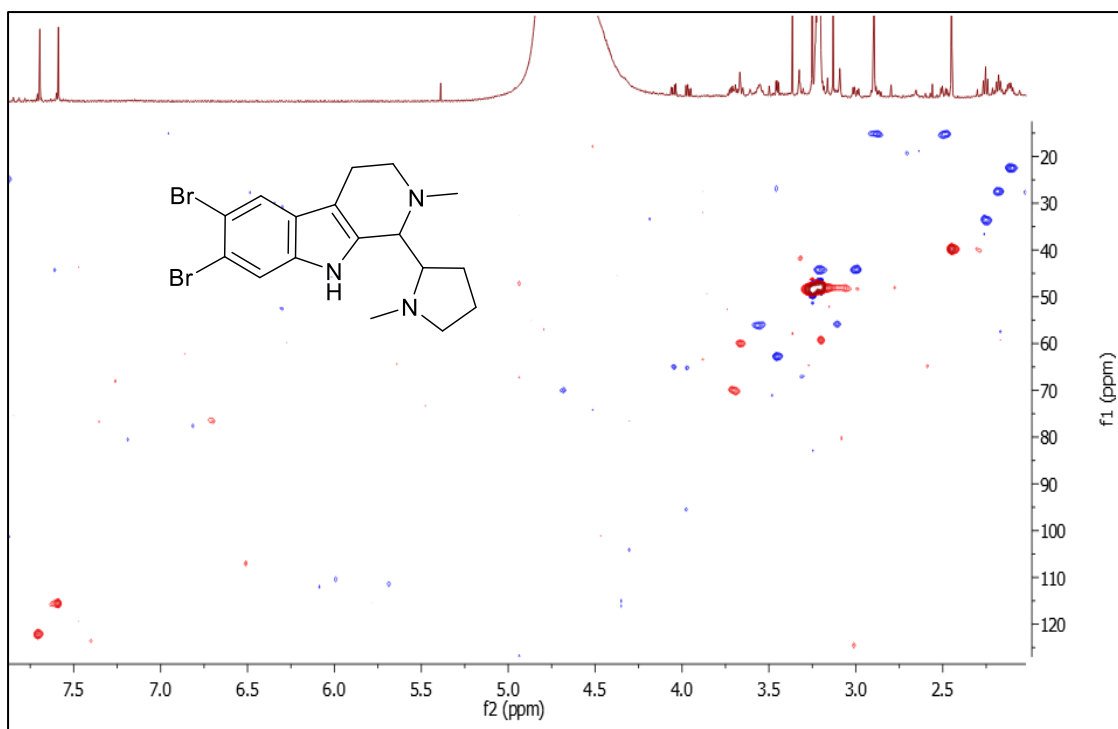


Figure S19. HSQC NMR (600 MHz, CD₃OD) spectrum of eudistomin Z2 (**2**)

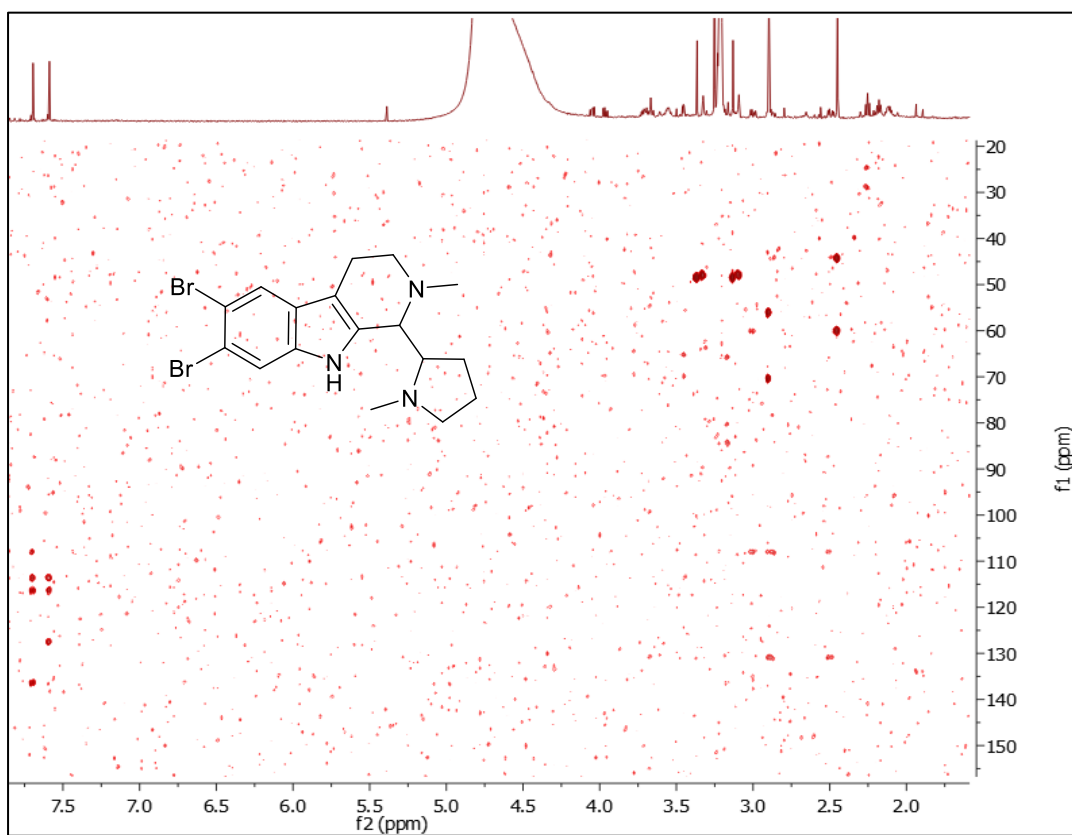


Figure S20. HMBC NMR (600 MHz, CD₃OD) spectrum of eudistomin Z2 (**2**)

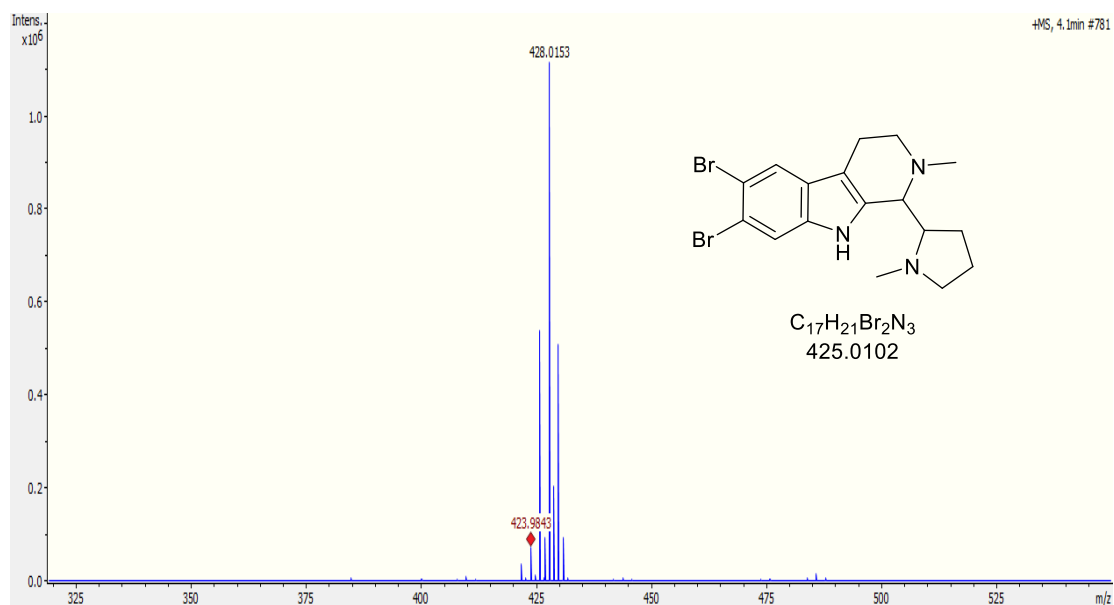


Figure S21. HRESIMS (ES^+) of eudistomin Z2 (**2**)

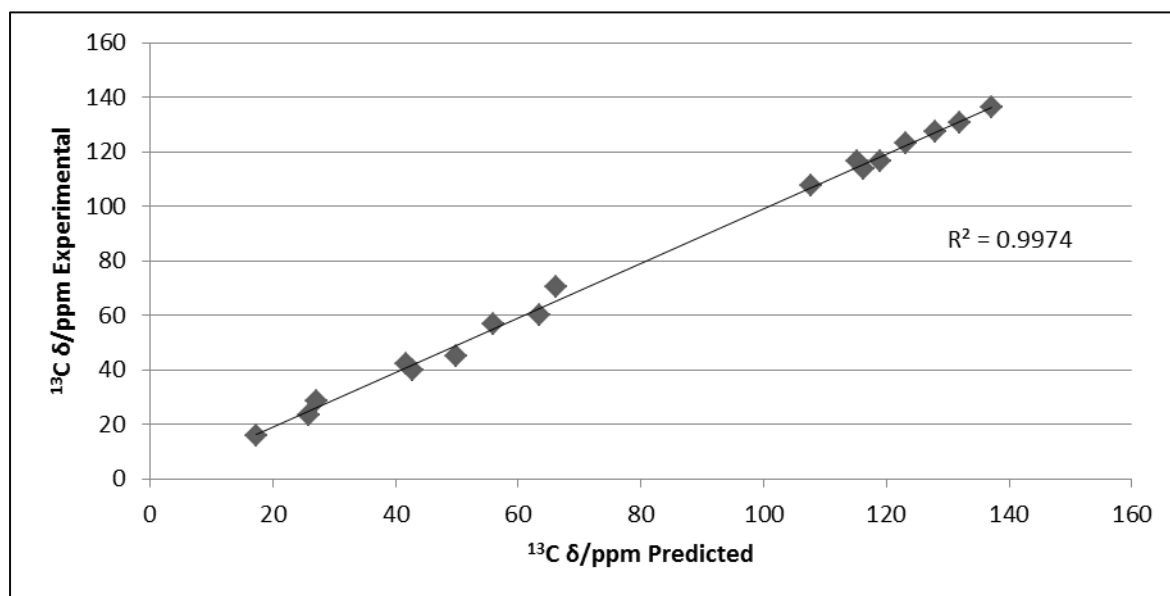


Figure S22. Predicted against experimental ^{13}C NMR data for eudistomin Z2 (**2**)

InputFile
C:\PharmaSea\output\IXS95.raw

MatrixDB
MatrixDB.txt

PharmaSeaDB
PharmaSeaDB.txt

Options
ApplyIX.mcr

FilterExpression
M+

threshold
200

logging_switch
true

rangeMass
0.0001

rangeRT
0.2

working
c:\PharmaSea\

MatrixTableName
▼

Create Query Rec
Matrix Append (Dynamic)
Generate Hits (Dynamic)

Run IX

Clear Output
Generate *MS1*
Generate *MS2*

Extracts MS1 for IX

OK Cancel

Figure S23. Screen shot showing icons used for calculating metrics: File name is loaded as an input file, then matrix database is selected then processing starts by clicking on creating a query record, followed by generation of HITS, clear output, and finally generation of MS1.

File	Edit	Format	View	Help
[16/05/2018 12:37:07]	Info:	typecsrcSpectrum		
[16/05/2018 12:37:08]	Info:	#Caption:TIC (ESI+)		
[16/05/2018 12:37:08]	Info:	22147483120		
[16/05/2018 12:37:11]	Info:	count:1		
[16/05/2018 12:37:11]	Info:	Before Count:0		
[16/05/2018 13:01:30]	Info:	After Count:681		
[16/05/2018 13:01:31]	Info:	Items in Matrix:1973		
[16/05/2018 13:01:32]	Info:	Items in PharmaSea:0		
[16/05/2018 13:01:32]	Info:	Index:0value:244.1146		
[16/05/2018 13:01:32]	Info:	Index:1value:774.4325		
[16/05/2018 13:01:32]	Info:	Index:2value:651.3972		
[16/05/2018 13:01:32]	Info:	Index:3value:686.3237		
[16/05/2018 13:01:32]	Info:	Index:4value:750.3585		
[16/05/2018 13:01:32]	Info:	Index:5value:210.137		
[16/05/2018 13:01:32]	Info:	Index:6value:997.5145		
[16/05/2018 13:01:32]	Info:	Index:7value:1139.5832		
[16/05/2018 13:01:32]	Info:	Index:8value:559.2675		
[16/05/2018 13:01:32]	Info:	Index:9value:538.3114		
[16/05/2018 13:01:32]	Info:	Index:10value:454.2569		
[16/05/2018 13:01:32]	Info:	Index:11value:1370.6875		
[16/05/2018 13:01:32]	Info:	Index:12value:712.4172		
[16/05/2018 13:01:32]	Info:	Index:13value:1121.5952		
[16/05/2018 13:01:32]	Info:	Index:14value:688.4267		
[16/05/2018 13:01:32]	Info:	Index:15value:904.4678		
[16/05/2018 13:01:32]	Info:	Index:16value:850.4357		
[16/05/2018 13:01:32]	Info:	Index:17value:481.2901		
[16/05/2018 13:01:32]	Info:	Index:18value:744.4817		
[16/05/2018 13:01:32]	Info:	Index:19value:742.4775		
[16/05/2018 13:01:32]	Info:	Index:20value:990.5395		
[16/05/2018 13:01:32]	Info:	Index:21value:608.3343		
[16/05/2018 13:01:32]	Info:	Index:22value:197.4084		
[16/05/2018 13:01:32]	Info:	Index:23value:314.0357		
[16/05/2018 13:01:32]	Info:	Index:24value:243.1228		
[16/05/2018 13:01:32]	Info:	Index:25value:1105.544		
[16/05/2018 13:01:32]	Info:	Index:26value:1114.6091		
[16/05/2018 13:01:32]	Info:	Index:27value:359.1864		
[16/05/2018 13:01:32]	Info:	Index:28value:278.1272		
[16/05/2018 13:01:32]	Info:	Index:29value:1017.5209		
[16/05/2018 13:01:32]	Info:	Index:30value:490.225		
[16/05/2018 13:01:32]	Info:	Index:31value:587.1464		
[16/05/2018 13:01:32]	Info:	Index:32value:754.3899		
[16/05/2018 13:01:32]	Info:	Index:33value:602.3032		
[16/05/2018 13:01:32]	Info:	Index:34value:602.3069		

Table S24. Some of the indexed masses after applying the scripted IX (IX.mcr).

RT:11.888906		
TIC Table		
No.	m/z	TIC(%)
3254	630.2781	0.51167
3255	630.2841	0.532145
3409	661.3071	0.624235
3410	661.3135	1.501414
3411	661.3199	2.277272
3412	661.3264	2.403202
3413	661.3328	1.783807
3414	661.3392	0.877303
3437	662.3168	0.627471
3438	662.3232	0.934815
3439	662.3297	0.973065
3440	662.3361	0.711841
3880	715.3231	0.730518
3881	715.3303	0.981028
3882	715.3376	0.926375
3883	715.3448	0.60829
4032	723.2294	0.965557
4033	723.2367	1.35508
4034	723.2441	1.332372
4035	723.2514	0.915089
4056	724.237	0.505375

Table S25. Example of masses that have been de-convoluted to yield m/z , retention time, and %TIC.

index	Notation	Adjusted PTM	PTM	RT	Area (counts)	Height (counts)
1	[M+H] ⁺ U ⁺ u	505.3498	506.3576	12.78968	7603551	983087.8
2	[M+H] ⁺ U ⁺ u	505.3498	506.3576	13.12518	5634764	726737
3	[M+H] ⁺ U ⁺ u	501.3157	502.3235	15.04968	3949936	536149.6
4	[M+H] ⁺ U ⁺ u	963.7047	964.7125	16.52152	924089.1	222148.7
5	[M+H] ⁺ U ⁺ u	1018.762	1019.77	18.82167	627349	128896
6	[M+H] ⁺ U ⁺ u	994.5527	995.5605	19.42801	479045.3	85171.1
7	[M+H] ⁺ U ⁺ u	970.388	971.3958	19.83518	813372	139478.6
8	[M+H] ⁺ U ⁺ u	447.2728	448.2806	19.96802	2117323	428831.5
9	[M+H] ⁺ U ⁺ u	1907.092	1908.1	20.12734	3653650	841166.1
10	[M+H] ⁺ U ⁺ u	942.5615	943.5693	20.15984	89983584	10703965
11	[M+Na] ⁺ U ⁺ u	942.5607	965.5505	20.19234	92330368	9061333
12	[M+H] ⁺ U ⁺ u	942.5602	943.568	20.28951	1.14E+08	10000632
13	[M+H] ⁺ U ⁺ u	964.5427	965.5505	20.32202	1.01E+08	8263899
14	[M+H] ⁺ U ⁺ u	682.1994	683.2072	20.51818	1163113	121939.5
15	[M+H] ⁺ U ⁺ u	1703.721	1704.729	20.61886	104236.6	24807.08
16	[M+H] ⁺ U ⁺ u	1789.766	1790.774	20.88434	132134.2	28332.2
17	[M+H] ⁺ U ⁺ u	648.5429	649.5508	20.95268	954145.4	130700.4

18	[M+H] ⁺ U ⁺ u	624.2548	625.2626	21.01818	1641136	295309.3
19	[M+H] ⁺ U ⁺ u	1875.812	1876.82	21.15468	151019.1	31739.14
20	[M+H] ⁺ U ⁺ u	636.2946	637.3024	21.22068	4937760	873317.4
21	[M+H] ⁺ U ⁺ u	480.3033	481.3111	21.28884	6329922	635843.3
22	[M+H] ⁺ U ⁺ u	622.2787	623.2865	21.32251	433505.3	101559.4
23	[M+H] ⁺ U ⁺ u	1960.839	1961.846	21.42535	119491.9	27328.34
24	[M+H] ⁺ U ⁺ u	638.2741	639.2819	21.75735	1681029	278350.8
25	[M+H] ⁺ U ⁺ u	622.2767	623.2846	21.89318	442831.5	115678.4
26	[M+H] ⁺ U ⁺ u	582.4971	583.5049	21.96034	801304.9	145924.9
27	[M+H] ⁺ U ⁺ u	372.2602	373.268	22.09285	10783159	1282189
28	[M+H] ⁺ U ⁺ u	1207.003	1208.011	22.39234	81264.33	23779.89
29	[M+H] ⁺ U ⁺ u	692.5679	693.5758	22.39234	2506138	310395.7
30	[M+H] ⁺ U ⁺ u	622.2793	623.2872	22.56052	1391955	207071
31	[M+H] ⁺ U ⁺ u	534.2605	535.2683	22.62751	12208508	2034984
32	[M+H] ⁺ U ⁺ u	650.2722	651.28	22.76317	1559739	148629.4
33	[M+H] ⁺ U ⁺ u	564.2347	565.2425	22.93001	1894161	226296.9
34	[M+H] ⁺ U ⁺ u	726.2787	727.2865	22.98334	1175189	180703.7
35	[M+Na] ⁺ U ⁺ u	380.2921	403.2819	23.01601	1976453	242713.6
36	[M+H] ⁺ U ⁺ u	380.2922	381.3	23.01601	25543202	3260617
37	[M+H] ⁺ U ⁺ u	622.2766	623.2844	23.21534	1042026	143729.9
38	[M+H] ⁺ U ⁺ u	964.5427	965.5506	23.34985	3691525	519056.5
39	[M+H] ⁺ U ⁺ u	964.5427	965.5506	23.58234	2768314	632021.2
40	[M+H] ⁺ U ⁺ u	650.5561	651.5639	23.75134	1666171	226210.4
41	[M+H] ⁺ U ⁺ u	636.2946	637.3024	23.85534	1797456	369957.4
42	[M+H] ⁺ U ⁺ u	658.274	659.2818	23.88901	2179307	299158.3
43	[M+H] ⁺ U ⁺ u	964.5427	965.5506	23.92368	4579825	563518.1
44	[M+H] ⁺ U ⁺ u	522.3167	523.3245	23.95834	47386972	1910936
45	[M+H] ⁺ U ⁺ u	282.2547	283.2625	24.26318	5038764	767560.3
46	[M+Na] ⁺ U ⁺ u	636.292	659.2818	24.63318	1619361	199234.2
47	[M+H] ⁺ U ⁺ u	636.2945	637.3024	24.63318	2022628	254647.4
48	[M+H] ⁺ U ⁺ u	650.3098	651.3176	24.73485	1353343	256827.3
49	[M+Na] ⁺ U ⁺ u	650.3096	673.2994	24.73485	7669456	1264348
50	[M+H] ⁺ U ⁺ u	600.2322	601.24	24.76868	951875.3	135099.3
51	[M+H] ⁺ U ⁺ u	605.4601	606.4679	24.83701	1871694	288427.2
52	[M+H] ⁺ U ⁺ u	310.2872	311.295	24.93801	2861254	800079

Table S26. Non-identified masses (not identified in either matrixDB or PharmaSeaDB) showing retention times, peak areas and peak heights based on ion counts.

No	Sample code	Sample complexity	Chemical novelty	Sample Media contamination	Sample novelty	Sample diversity
1	Tavarua-8-100	69	0.07	6.00E-02	0.93	2.53E+09
2	Eden-6-50	109	0.11	6.00E-03	0.89	3.31E+09
3	Eden-6-25	68	0.13	4.00E-03	0.87	9.67E+07
4	Balolo-3.14-100+TFA	92	0.17	8.66E-02	0.83	2.70E+08
5	Eden-6-100	135	0.18	1.26E-02	0.82	3.39E+09
6	Tavaniko-4-100+TFA	75	0.24	9.74E-03	0.76	9.44E+08
7	Namotu-2-100+TFA	70	0.24	9.00E-03	0.76	1.47E+09
8	Lau-1-25	70	0.24	9.00E-03	0.76	1.47E+09
9	Pinacple-5-100+TFA	23	0.26	3.00E-04	0.74	1.67E+08
10	Buago-3-25	30	0.27	4.33E-03	0.73	5.82E+08
11	Namotu-2-100	70	0.27	1.00E-02	0.73	2.26E+09
12	Tavaniko-7-100+TFA	62	0.27	9.00E-03	0.73	1.17E+08
13	Tavaniko-5-25	25	0.28	4.00E-03	0.72	3.50E+08
14	Tavaniko-7-100	67	0.28	1.00E-02	0.72	1.10E+08
15	Tavaniko-4-100	95	0.33	1.68E-02	0.67	5.97E+08
16	Pinacple-5-100	140	0.33	2.49E-02	0.67	9.54E+09
17	Namotu- 2-25	15	0.33	2.60E-03	0.67	1.14E+08
18	Tavaniko-4-25	20	0.35	3.79E-03	0.65	5.43E+08
19	Tavaniko-5-100	94	0.37	1.89E-02	0.63	2.06E+09
20	Tavaniko-5-100+TFA	94	0.37	1.89E-02	0.63	2.06E+08
21	Buago-7-100	126	0.39	2.65E-02	0.61	9.76E+09
22	Tavaniko-7-25	23	0.39	4.00E-03	0.61	1.67E+07
23	Balolo-3.14-25	35	0.40	7.50E-03	0.60	1.62E+09
24	Balolo-3.14-100	69	0.41	1.50E-02	0.59	2.17E+09
25	Eden-4-100+TFA	33	0.42	7.36E-03	0.58	5.41E+08
26	Eden-5-100+TFA	62	0.47	1.50E-02	0.53	7.93E+08
27	Eden-4-100	25	0.48	6.30E-03	0.52	5.15E+07
28	Tavarua-8-100+TFA	111	0.49	2.93E-02	0.51	3.23E+09
29	Balolo-3.14-50	53	0.51	1.46E-02	0.49	2.95E+09
30	Buago-7-100+TFA	113	0.52	3.19E-02	0.48	3.13E+09
31	Eden-6-100+TFA	108	0.52	3.00E-02	0.47	2.06E+09
32	Lau-1-100+TFA	96	0.54	4.00E-02	0.46	8.14E+08
33	Tavarua-2-25	63	0.56	1.89E-02	0.44	1.22E+09
34	Eden-4-25	18	0.56	5.30E-03	0.44	1.34E+08
35	Buago-3-100	88	0.56	2.65E-02	0.44	1.38E+08
36	Pinacple-5-25	58	0.52	1.60E-03	0.43	1.21E+09
37	Buago-3-50	56	0.59	1.79E-02	0.41	1.17E+09
38	Namotu-2-50	30	0.60	9.00E-03	0.40	5.49E+09
39	Tavarua-2-100+TFA	153	0.61	5.10E-02	0.39	6.15E+09
40	Pinacple-5-50	96	0.61	3.20E-02	0.39	7.97E+09
41	Tavaniko-4-50	47	0.67	1.57E-02	0.38	6.86E+08
42	Eden-5-25	57	0.63	1.90E-02	0.37	1.10E+09
43	Eden-5-100	30	0.63	1.00E-02	0.37	6.70E+08
44	Tavarua-2-100	161	0.63	5.50E-02	0.37	6.90E+09
45	Buago-3-100+TFA	69	0.64	2.38E-02	0.36	1.09E+09
46	Eden-4-50	36	0.64	1.21E-02	0.36	7.44E+08
47	Eden-5-50	64	0.64	2.20E-02	0.36	2.55E+09
48	Tavaniko-7-50	48	0.65	1.63E-02	0.35	9.03E+08
49	Tavarua-8-50	108	0.66	3.73E-02	0.34	2.71E+09
50	Buago-7-25	41	0.66	1.46E-02	0.34	7.55E+07
51	Tavarua-8-25	66	0.66	5.50E-02	0.33	1.19E+09
52	Buago-7-50	67	0.67	2.44E-02	0.33	2.07E+09
53	Tavaniko-5-50	68	0.68	2.50E-02	0.32	2.21E+09
54	Lau-1-100	79	0.72	3.08E-02	0.28	9.08E+08
55	Lau-1-50	79	0.76	3.20E-02	0.24	9.08E+09
56	Tavarua-2-50	74	0.66	3.30E-02	0.15	2.52E+09

Table S27. Metric data showing 'Sample novelty' arranged in decreasing order.

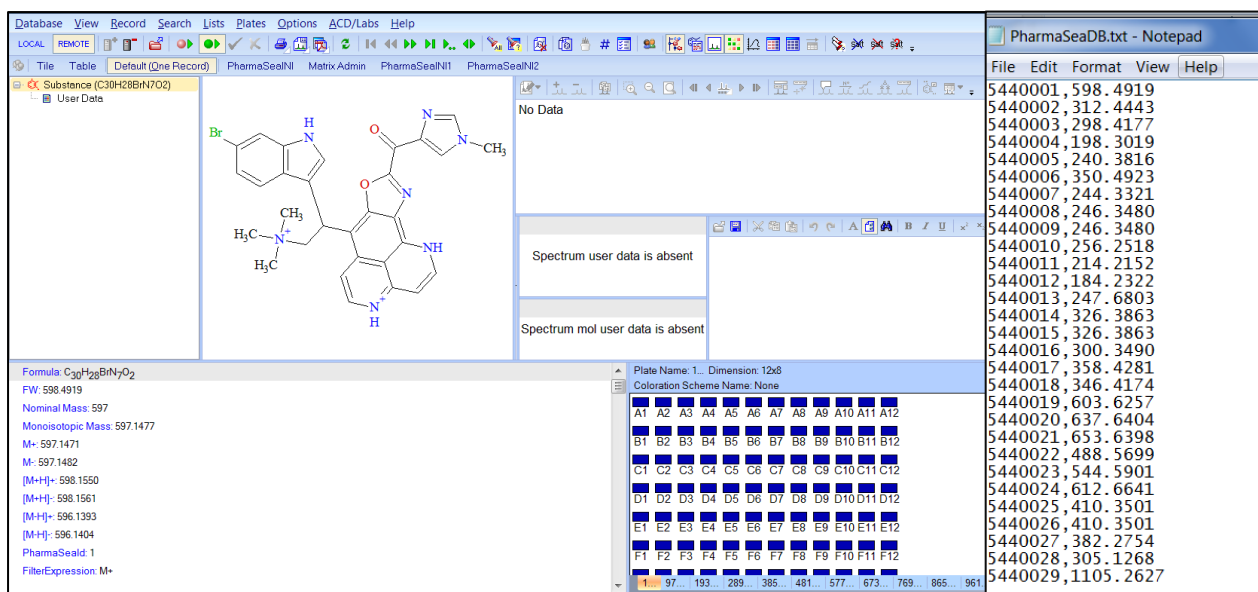


Figure S28. Screen shot of PharmaSeaDB with some of the 26,618 masses (notepad text file).

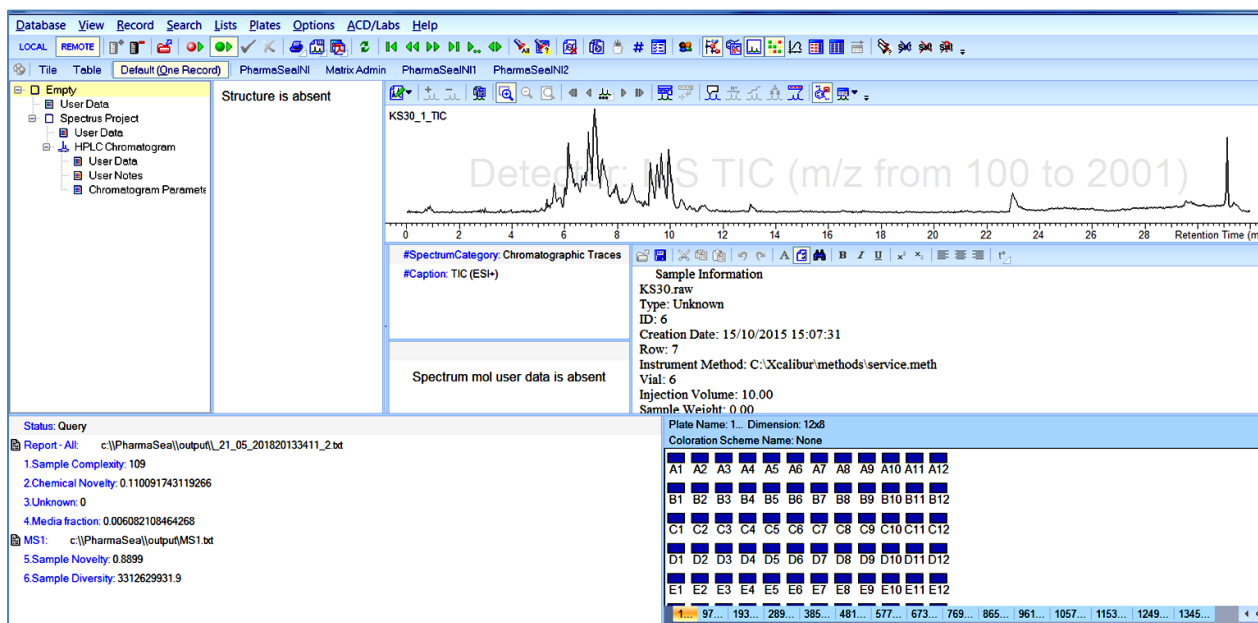


Figure S29. Showing metric data for eden-6-SPE-50 in PharmaSeaDB.

[21/05/2018 20:34:35]	Info:	Row:1634data:743.563902RT24.546125
[21/05/2018 20:34:35]	Info:	Found:743.563902
[21/05/2018 20:34:35]	Info:	Row:1656data:754.782915RT12.780797
[21/05/2018 20:34:35]	Info:	No Matrix data found
[21/05/2018 20:34:35]	Info:	No PharmaSea data found
[21/05/2018 20:34:35]	Info:	Row:1665data:761.829413RT9.281131
[21/05/2018 20:34:35]	Info:	Found:761.829413
[21/05/2018 20:34:35]	Info:	Row:1689data:776.844150RT9.446795
[21/05/2018 20:34:35]	Info:	No Matrix data found
[21/05/2018 20:34:35]	Info:	No PharmaSea data found
[21/05/2018 20:34:35]	Info:	Row:1708data:787.861499RT9.764294
[21/05/2018 20:34:35]	Info:	No Matrix data found
[21/05/2018 20:34:35]	Info:	No PharmaSea data found
[21/05/2018 20:34:35]	Info:	Row:1718data:803.880017RT9.699630
[21/05/2018 20:34:35]	Info:	No Matrix data found
[21/05/2018 20:34:35]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1741data:833.850983RT9.414291
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1752data:936.877261RT9.513792
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1756data:940.876052RT10.096458
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found

[21/05/2018 20:34:36]	Info:	Row:1771data:958.854090RT9.513792
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1796data:980.870398RT7.848295
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1808data:990.902145RT8.449133
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1829data:1012.892286RT8.548959
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1856data:1023.655738RT23.302458
[21/05/2018 20:34:36]	Info:	Found:1023.655738
[21/05/2018 20:34:36]	Info:	Row:1882data:1063.925764RT6.873798
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1921data:1176.953009RT6.709305
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:36]	Info:	No PharmaSea data found
[21/05/2018 20:34:36]	Info:	Row:1924data:1177.960116RT7.003458
[21/05/2018 20:34:36]	Info:	No Matrix data found
[21/05/2018 20:34:37]	Info:	No PharmaSea data found
[21/05/2018 20:34:37]	Info:	Row:1981data:1323.577000RT7.099458
[21/05/2018 20:34:37]	Info:	No Matrix data found

[21/05/2018 20:34:37]	Info:	No PharmaSea data found
[21/05/2018 20:34:37]	Info:	New entity count97

Table S30. Part of the database search output file for eden-6-SPE-50 showing some of the masses found or not found in PharmaSeaDB (found masses in yellow).

index	Notation	Adjusted PTM	PTM	RT	Area (counts)	Height (counts)
1	[M+H] ⁺ U ⁺ u	654.042575	655.050395	5.091125	1722767.625	325918.7188
2	[M+H] ⁺ U ⁺ u	644.051817	645.059637	5.190792	524280.7188	107164.7734
3	[M+H] ⁺ U ⁺ u	355.533319	356.541139	5.323465	370685.4063	107726.5313
4	[M+H] ⁺ U ⁺ u	547.039235	548.047055	5.35696	3366301.75	774876.1875
5	[M+H] ⁺ U ⁺ u	654.042575	655.050395	5.389463	2586024.5	462463.9375
6	[M+H] ⁺ U ⁺ u	725.080014	726.087834	5.456626	1508335	333517.2813
7	[M+H] ⁺ U ⁺ u	697.067274	698.075094	5.524802	885161.5	167362.4375
8	[2M+H] ⁺ U ⁺ u	355.5339465	712.075713	5.591292	2247208.75	516536.9688
9	[M+H] ⁺ U ⁺ u	355.533617	356.541437	5.591292	833859.25	202181.6875
10	[M+H] ⁺ U ⁺ u	392.976768	393.984588	5.657458	317917.5313	74997.10938
11	[M+H] ⁺ U ⁺ u	387.976938	388.984758	5.657458	5071848	472594.7813
12	[M+H] ⁺ U ⁺ u	427.048119	428.055939	5.861292	1382485.25	327103.7813
13	[M+H] ⁺ U ⁺ u	294.87535	295.88317	5.894797	55263.45313	18421.15039
14	[M+H] ⁺ U ⁺ u	481.002037	482.009857	5.929795	608498.8125	156529.1406
15	[M+H] ⁺ U ⁺ u	411.05316	412.06098	5.997968	3447804.5	677845.375
16	[M+H] ⁺ U ⁺ u	618.024351	619.032171	6.096625	514171.8438	89936.88281
17	[M+H] ⁺ U ⁺ u	137.057956	138.065776	6.17813	23029454	3971565.5
18	[M+H] ⁺ U ⁺ u	657.986311	658.994131	6.24413	7250306	973462.9375
19	[M+Na] ⁺ U ⁺ u	657.987422	680.977192	6.276633	456254.1875	60125.34375
20	[M+H] ⁺ U ⁺ u	516.95908	517.9669	6.477792	12232077	2384149.25
21	[M+Na] ⁺ U ⁺ u	517.001034	539.990804	6.5113	1323050.875	208922.0313
22	[M+H] ⁺ U ⁺ u	389.975316	390.983136	6.544792	2999039	588073.5625
23	[M+H] ⁺ U ⁺ u	657.986308	658.994128	6.642295	33492938	4386796.5
24	[M+Na] ⁺ U ⁺ u	657.983726	680.973496	6.642295	2712141.25	399185.5
25	[M+H] ⁺ U ⁺ u	293.870605	294.878425	6.709305	16146086	1297266.125
26	[M+H] ⁺ U ⁺ u	1175.945189	1176.953009	6.709305	60042.23828	20014.07813
27	[M+H] ⁺ U ⁺ u	516.958996	517.966816	6.742795	11739371	2095622.75
28	[M+H] ⁺ U ⁺ u	492.995693	494.003513	6.840292	17253944	3352029.25
29	[M+H] ⁺ U ⁺ u	1062.917944	1063.925764	6.873798	440745.5313	115403.8594
30	[M+H] ⁺ U ⁺ u	123.042618	124.050438	6.937292	47897256	7130761.5
31	[M+H] ⁺ U ⁺ u	548.974903	549.982723	7.003458	28748780	4735454
32	[M+H] ⁺ U ⁺ u	1176.952296	1177.960116	7.003458	61022.73438	20340.91211
33	[M+H] ⁺ U ⁺ u	1322.56918	1323.577	7.099458	495205.2188	137363.5938
34	[M+Na] ⁺ U ⁺ u	529.001644	551.991414	7.157458	133046856	21188410
35	[M+H] ⁺ U ⁺ u	528.958181	529.966001	7.157458	7593552	1071769.25
36	[M+H] ⁺ U ⁺ u	295.869788	296.877608	7.189958	6543074	567664.875
37	[M+H] ⁺ U ⁺ u	520.969052	521.976872	7.22247	74022168	10566150

38	[M+H] ⁺ U ⁺ u	532.971645	533.979465	7.288468	11563117	1761519
39	[M+H] ⁺ U ⁺ u	509.009537	510.017357	7.355627	4404338	883548.375
40	[M+H] ⁺ U ⁺ u	593.006551	594.014371	7.38913	5709940.5	547367.375
41	[M+H] ⁺ U ⁺ u	566.979358	567.987178	7.38913	5589680	852864.5625
42	[M+H] ⁺ U ⁺ u	133.063447	134.071267	7.454625	32177950	6051857.5
43	[M+H] ⁺ U ⁺ u	293.870735	294.878555	7.488305	50051172	1720277.75
44	[M+H] ⁺ U ⁺ u	580.993782	582.001602	7.585793	61691980	9076847
45	[M+H] ⁺ U ⁺ u	564.996708	566.004528	7.650793	10564537	1705035.75
46	[M+H] ⁺ U ⁺ u	592.991328	593.999148	7.715958	2396605.25	795590.5
47	[M+H] ⁺ U ⁺ u	590.995016	592.002836	7.749625	5151127.5	364687.875
48	[M+H] ⁺ U ⁺ u	587.000267	588.008087	7.782125	995906.6875	139206.1094
49	[2M+H] ⁺ U ⁺ u	489.931289	980.870398	7.848295	273027.2188	65247.83594
50	[M+H] ⁺ U ⁺ u	316.868408	317.876228	7.8808	5533428	300760.625
51	[M+H] ⁺ U ⁺ u	480.008273	481.016093	7.980625	7888932.5	1550126.25
52	[M+H] ⁺ U ⁺ u	580.993294	582.001114	8.046125	9656500	1948285.5
53	[M+Na] ⁺ U ⁺ u	619.011468	642.001238	8.217463	225438.7188	41064.64844
54	[M+H] ⁺ U ⁺ u	619.024842	620.032662	8.28246	751408.5625	119020.6641
55	[M+H] ⁺ U ⁺ u	668.954974	669.962794	8.317133	800510.4375	144649.9219
56	[M+H] ⁺ U ⁺ u	346.861762	347.869582	8.382125	2575860.5	526560.75
57	[M+H] ⁺ U ⁺ u	989.894325	990.902145	8.449133	625771.375	100623.9688
58	[M+H] ⁺ U ⁺ u	293.870607	294.878427	8.516292	16406955	1154891.25
59	[M+Na] ⁺ U ⁺ u	989.902516	1012.892286	8.548959	228294.3594	54543.42969
60	[M+H] ⁺ U ⁺ u	318.866417	319.874237	8.614959	6111329.5	418850.8438
61	[M+H] ⁺ U ⁺ u	662.921088	663.928908	8.614959	4451139.5	900074.0625
62	[M+H] ⁺ U ⁺ u	564.996751	566.004571	8.679967	1142288.75	213335.5
63	[M+H] ⁺ U ⁺ u	608.989049	609.996869	8.712459	2974877	590011.5
64	[M+H] ⁺ U ⁺ u	690.949434	691.957254	8.812134	1518258	279666.3125
65	[M+H] ⁺ U ⁺ u	657.986369	658.994189	8.880458	8979836	1117340.25
66	[M+H] ⁺ U ⁺ u	609.023162	610.030982	9.013636	360388	61931.31641
67	[M+Na] ⁺ U ⁺ u	654.013574	677.003344	9.079803	108453.9453	18757.08398
68	[M+H] ⁺ U ⁺ u	654.010308	655.018128	9.079803	1528109.375	311130.6875
69	[M+H] ⁺ U ⁺ u	670.001028	671.008848	9.348125	2190130.25	415824.3438
70	[M+H] ⁺ U ⁺ u	832.843163	833.850983	9.414291	582660.4375	120944.4297
71	[M+H] ⁺ U ⁺ u	775.83633	776.84415	9.446795	1021333.375	233024.0469
72	[M+H] ⁺ U ⁺ u	935.869441	936.877261	9.513792	1028585.313	244727.6875
73	[M+Na] ⁺ U ⁺ u	935.86432	958.85409	9.513792	2518629.5	489982.625
74	[M+H] ⁺ U ⁺ u	518.955086	519.962906	9.546291	7491887	1527384.75
75	[M+H] ⁺ U ⁺ u	295.869212	296.877032	9.546291	8312756.5	1085601.125
76	[M+H] ⁺ U ⁺ u	802.872197	803.880017	9.69963	102525512	17427584
77	[M+H] ⁺ U ⁺ u	786.853679	787.861499	9.764294	1240423.625	171824.0781
78	[M+H] ⁺ U ⁺ u	449.941058	450.948878	9.893958	19404142	2001018.25
79	[M+H] ⁺ U ⁺ u	479.952502	480.960322	9.961125	45008736	7225933.5
80	[M+H+NH ⁺ D3 ⁺ d] ⁺ U ⁺ u	531.943254	532.951074	10.029469	9509643	1432760
81	[M+Na] ⁺ U ⁺ u	514.959357	537.949127	10.029469	1515117.125	218289.0469
82	[M+H] ⁺ U ⁺ u	514.96125	515.96907	10.062961	958420	351743
83	[M+H] ⁺ U ⁺ u	939.868232	940.876052	10.096458	149508.7188	26675.91406

84	[M+H] ⁺ U+ ⁺ u	334.896974	335.904794	10.195298	2672096.25	428876.5938
85	[M+H] ⁺ U+ ⁺ u	718.98502	719.99284	10.364458	1638641.375	318444.375
86	[M+H] ⁺ U+ ⁺ u	646.963851	647.971671	10.466125	4086077.5	714921.375
87	[M+H] ⁺ U+ ⁺ u	316.868406	317.876226	10.566635	467375.625	112081.9297
88	[M+H] ⁺ U+ ⁺ u	318.866413	319.874233	10.63496	1091282.5	175160.2656
89	[M+H] ⁺ U+ ⁺ u	437.921711	438.929531	10.76896	7783699	1503519.125
90	[M+Na] ⁺ U+ ⁺ u	437.923212	460.912982	10.76896	1464272.75	236470.1563
91	[M+H] ⁺ U+ ⁺ u	451.9188	452.92662	11.139459	3554702.5	758443.4375
92	[M+H] ⁺ U+ ⁺ u	317.870448	318.878268	11.206625	17947542	2278856.25
93	[M+Na] ⁺ U+ ⁺ u	432.947698	455.937468	11.206625	2280204.5	287015.5313
94	[M+H] ⁺ U+ ⁺ u	432.95225	433.96007	11.240125	4905713	719683.875
95	[M+H] ⁺ U+ ⁺ u	753.775095	754.782915	12.780797	213109.4375	58364.19922
96	[M+H] ⁺ U+ ⁺ u	726.278664	727.286484	22.998459	1282153.875	193823
97	[M+H] ⁺ U+ ⁺ u	523.317427	524.325247	24.683792	7699646	604005.9375

Table S31. Masses not identified in PhamaSeaDB for eden-6-SPE-50.

index	Notation	Adjusted PTM	PTM	RT	Area (counts)	Height (counts)
1	[M+H] ⁺ U+ ⁺ u	289.142969	290.150789	3.92395	1863052.625	476671.125
2	[M+H] ⁺ U+ ⁺ u	489.084319	490.092139	4.75029	1419154.625	299124.75
3	[M+H] ⁺ U+ ⁺ u	581.200219	582.208039	5.017783	1961837.625	203842.9688
4	[M+H] ⁺ U+ ⁺ u	747.203381	748.211201	5.321126	531048.75	145307.5156
5	[M+H] ⁺ U+ ⁺ u	485.254819	486.262639	8.758117	1677232.375	383905.0625
6	[M+H] ⁺ U+ ⁺ u	606.256328	607.264148	10.872622	405220.4375	89519.46094
7	[M+H] ⁺ U+ ⁺ u	509.280301	510.288121	11.852622	2841848	507860.5625
8	[M+H] ⁺ U+ ⁺ u	682.19892	683.20674	20.509951	1118910.875	121562.4375
9	[M+H] ⁺ U+ ⁺ u	636.294562	637.302382	21.252451	1111424.75	174955.8594
10	[M+H] ⁺ U+ ⁺ u	670.500297	671.508117	22.745285	2481364.75	280006.6563
11	[M+H] ⁺ U+ ⁺ u	670.500297	671.508117	24.899786	284310.1563	56117.85547

Table S32. Masses not identified in PharmaSeaDB for Tavarua-2-SPE-50.

[05/06/2018 16:21:49]	Info:	No Matrix data found
[05/06/2018 16:21:49]	Info:	No PharmaSea data found
[05/06/2018 16:21:49]	Info:	Row:303data:680.429534RT13.956450
[05/06/2018 16:21:49]	Info:	Found:680.429534
[05/06/2018 16:21:49]	Info:	Row:307data:683.206740RT20.509951
[05/06/2018 16:21:49]	Info:	No Matrix data found
[05/06/2018 16:21:49]	Info:	No PharmaSea data found

[05/06/2018 16:21:49]	Info:	Row:337data:743.563902RT24.561449
[05/06/2018 16:21:49]	Info:	Found:743.563902
[05/06/2018 16:21:49]	Info:	Row:342data:748.211201RT5.321126
[05/06/2018 16:21:50]	Info:	No Matrix data found
[05/06/2018 16:21:50]	Info:	No PharmaSea data found
[05/06/2018 16:21:50]	Info:	Row:345data:755.279947RT6.422616
[05/06/2018 16:21:50]	Info:	No Matrix data found
[05/06/2018 16:21:50]	Info:	No PharmaSea data found
[05/06/2018 16:21:50]	Info:	Row:348data:761.845353RT10.607117
[05/06/2018 16:21:50]	Info:	Found:761.845353
[05/06/2018 16:21:50]	Info:	Row:370data:825.232955RT5.353617
[05/06/2018 16:21:50]	Info:	Found:825.232955
[05/06/2018 16:21:50]	Info:	Row:372data:833.869207RT10.806622
[05/06/2018 16:21:50]	Info:	Found:833.869207
[05/06/2018 16:21:50]	Info:	Row:402data:1015.258717RT8.190117
[05/06/2018 16:21:50]	Info:	No Matrix data found
[05/06/2018 16:21:50]	Info:	No PharmaSea data found
[05/06/2018 16:21:50]	Info:	Row:422data:1043.293635RT8.625283
[05/06/2018 16:21:50]	Info:	No Matrix data found
[05/06/2018 16:21:50]	Info:	No PharmaSea data found
[05/06/2018 16:21:50]	Info:	Row:459data:1775.699654RT7.783950
[05/06/2018 16:21:50]	Info:	Found:1775.699654
[05/06/2018 16:21:50]	Info:	New entity count25

Table S33. A selection of masses identified and not identified in PharmaSeaDB for tavarua-2-SPE-50

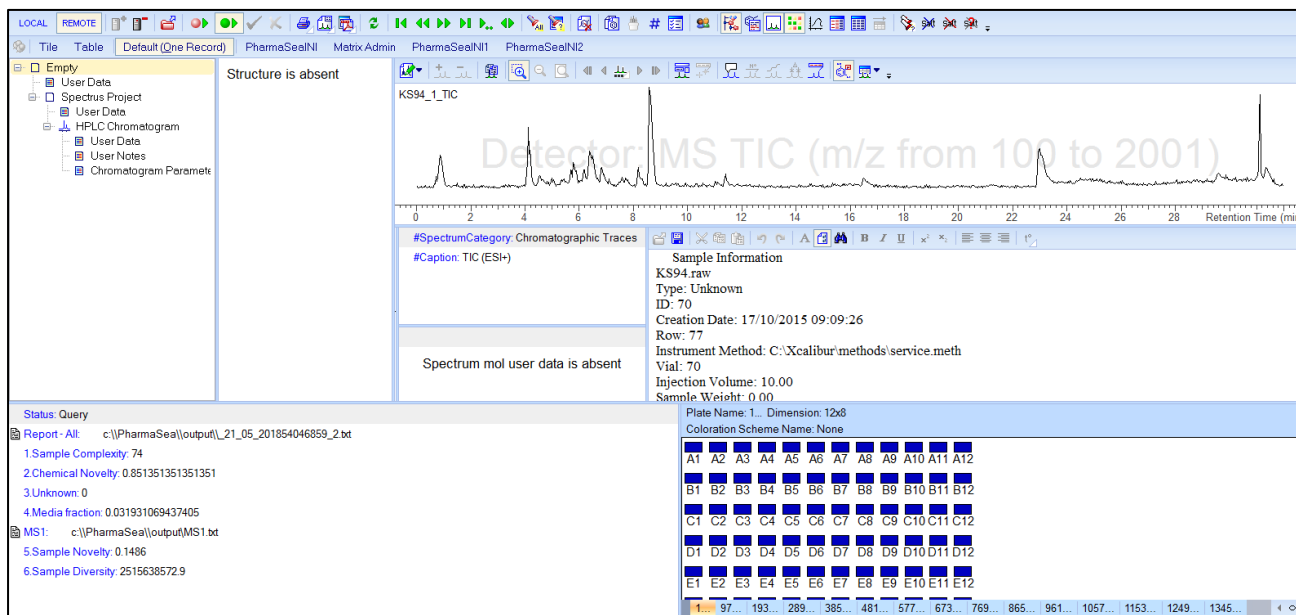


Figure S34. Metric data for tavarua-2-SPE-50 in PharmaSeaDB.

DATA PROCESSING PARAMETERS

[Component Window]

Parameter=% Peak width FWHM

Value=30

FWHM Deviation (%)=50

[Mass Value Window (Da)]

Value=0.5

Mass Values=Accurate

Type of Mass=Peak Top

[12C/13C Error (%)]

Value=10

[Ionization]

Type=Auto

[Polarity]

Auto=True

Positive=True

[Mass Accuracy]

Units=Da

Value=0.5

MC Display (Da)=0.5

MC Calculation (Da)=1

m/z Values=Averaged

Mass Unity Range Shift=0.2

[CODA]

Apply=True
Peak Minimum FWHM=Scans
Value=1
Peak Minimum Height (% of max)=0
Peak S/N Threshold=100
Peak Area Threshold (counts)=100000
Pick the Number of Most Abundant Peaks=False
Number of Peaks=1
Apply Smoothing=True
Method=Savitsky-Golay
Shape of FFT=Fix
Fix V1(%)=5
Fix V2(%)=50
Polynomial=1
Points=3
Enhancement=True
Baseline Correction=False
Box Half Width (Scans)=5
Noise Factor=5
MCQ Value=1
Abundance Value (% of TIC)=0

[Components Analysis]

Apply=True
Default Peak Width=0.05
Default Peak Function=Gauss
Lorentz Fraction=1
Vary position within=0.1
Max width limit=10
Min width limit=0.001
Tolerance=3
Max number of iteration=150
Lock Width=True
Lock Height=True
Fixed Width=False
Parameter=0.05
Max Number of Peaks per MC=15
Max Number of Co-eluted Peaks per Region=15

[Missing Isotopes Search]

Apply=True
Good=True
Poor=True
Review=True
Depth of Analysis=3

[Area Threshold Analysis]

Apply=True
Tangent Slope=0.05

[FWHM Threshold Analysis]

Apply=False
Tangent Slope=0.05

[Height Threshold Analysis]

Apply=False
Tangent Slope=0.05

[Search for Adducts]

Proton=1
Ammonia=0
Sodium=1
Potassium=0
Acetonitrile=0
Methanol=0
H2O=0
Lithium=0
Acetate=1
Formate=1
TFA=1

[Search for Multimers]

Proton=2
Ammonia=0
Sodium=1
Potassium=0
Acetonitrile=0
Methanol=0
H2O=0
Lithium=0
Acetate=2
Formate=2
TFA=2

[A + 2]

Percent=30
Abundance Ratio Error (%)=50
Sulfur=True
Sulfur Min=0
Sulfur Max=3
Chlorine=True
Chlorine Min=0
Chlorine Max=3
Bromine=True
Bromine Min=0
Bromine Max=3

[M Ion Error (%)]

Value=20

[Additional M Ion Correction (%)]

Value=5

[Compare Options]

tR Window (min)=0.2

Abundance=Area

Abundance Ratio=5:1

Isotope Cluster Tolerance=True

Centre of Gravity (%)=50

Cluster Dispersion (%)=50

Boolean Comparison=False

[Keep the following Peaks only]

Notated as Molecular Ions=True

Peaks with other textual descriptors=True

Classified as Good=True

Classified as Review=True

Classified as Poor=True

[Apply Ion Presence Detection for]

Apply=True

Molecular Ions=True

Adduct Ions=True

Primary Fragment Ions=True

Multimer Ions=False

Unconfirmed 12C Ions=False

Unconfirmed 13C Ions=False

Unassigned Ions=False

Label Parent Ion Modifications=False

[Apply Ion Presence Detection Threshold]

Apply=True

Abundance Type=Height

Abundance Units=%

Value=1

[Apply Ion Presence Detection 12C/13C Filter]

Apply=False

12C/13C (Pass)=True

12C/13C (Fail)=False

12C/13C (None)=False

[Label Possible Metabolites]

Apply=False

[Peak Quality Analysis]

W (Area)=20

W (Height)=20

W (Asymmetry)=10

W (S/N)=0

W (Carbon)=20

W (12C/13C)=20

W (Class)=10

W (Uniqueness)=0

[Calculate the Following Mass Values]

Peak Top=True

Peak Top Window Type=% of Peak FWHM

Peak Top Window Value=100

Left Shoulder=False

Exclude the First Scan=True

Right Shoulder=False

Exclude the Last Scan=True

Entire Peak=False

Exclude the First and Last Scans=True

Apply Intensity Based Weighting=True

[Data Analysis Region]

Full=False

Start tR=1

End tR=19

[Structure Search]

Apply=False

Mass Values=Accurate

Mass Accuracy=0.001

Type of Mass=Peak Top

Molecular Ions=True

¹²C Ions=False

Unassigned Ions=False

Directory=

Label (User Data field)=

¹²C/¹³C (Pass)=True

¹²C/¹³C (Fail)=True

¹²C/¹³C (None)=True

[Auto Neutral Loss Determination]

Apply=False

Display Accuracy (Decimal Places)=2

[Convert Profile to Centroid]

Apply=False

Mass Values to Center=Top

Abundance to Center=Area

Peak Picking Minimum FWHM (Points)=3

Abundance Threshold (%)=1

Centroid %=80

Apply Smoothing=False

Polynomial=1

Window Width (Points)=3

[Component MC Summing]

Apply=False

[S/N Calculation]

Parameter=Scans

Scans=50

tR=0.1

[Mass Accuracy for Peak Labeling]

Type of Mass=Peak Top

Mass Accuracy=0.0025

[HRMS Data Processing Approach]

Apply=True

Type of Mass=Peak Top

Mass Value Window (Da, +/-)=0.0025

[Extract Pure Component Spectra]

Apply=False

Extract Selected Components Only=True

Directory=

File Name=

Action if File Name Exists=AutoName

[Peak Quality Threshold]

Apply=False

Value=1

Figure S35. Parameters used by IntelliXtrac (IX.mcr) for processing of LC-HRMS data. These are the optimised data processing settings used to provide greater mass accuracy of the detected ions.

No	Sample code	Sample complexity	Chemical novelty	Sample Media contamination	Sample novelty	Sample diversity
1	Tavarua-2-100	161	0.63	5.50E-02	0.37	6.90E+09
2	Tavarua-2-100+TFA	153	0.61	5.10E-02	0.39	6.15E+09
3	Pinacple-5-100	140	0.33	2.49E-02	0.67	9.54E+09
4	Eden-6-100	135	0.18	1.26E-02	0.82	3.39E+09
5	Buago-7-100	126	0.39	2.65E-02	0.61	9.76E+09
6	Buago-7-100+TFA	113	0.52	3.19E-02	0.48	3.13E+09
7	Tavarua-8-100+TFA	111	0.49	2.93E-02	0.51	3.23E+09
8	Eden-6-50	109	0.11	6.00E-03	0.89	3.31E+09
9	Eden-6-100+TFA	108	0.52	3.00E-02	0.47	2.06E+09
10	Tavarua-8-50	108	0.66	3.73E-02	0.34	2.71E+09
11	Lau-1-100+TFA	96	0.54	4.00E-02	0.46	8.14E+08
12	Pinacple-5-50	96	0.61	3.20E-02	0.39	7.97E+09
13	Tavaniko-4-100	95	0.33	1.68E-02	0.67	5.97E+08
14	Tavaniko-5-100	94	0.37	1.89E-02	0.63	2.06E+09
15	Tavaniko-5-100+TFA	94	0.37	1.89E-02	0.63	2.06E+08
16	Balolo-3.14-100+TFA	92	0.17	8.66E-02	0.83	2.70E+08
17	Buago-3-100	88	0.56	2.65E-02	0.44	1.38E+08
18	Lau-1-100	79	0.72	3.08E-02	0.28	9.08E+08
19	Lau-1-50	79	0.76	3.20E-02	0.24	9.08E+09
20	Tavaniko-4-100+TFA	75	0.24	9.74E-03	0.76	9.44E+08
21	Tavarua-2-50	74	0.66	3.30E-02	0.15	2.52E+09
22	Namotu-2-100+TFA	70	0.24	9.00E-03	0.76	1.47E+09
23	Lau-1-25	70	0.24	9.00E-03	0.76	1.47E+09
24	Namotu-2-100	70	0.27	1.00E-02	0.73	2.26E+09
25	Tavarua-8-100	69	0.07	6.00E-02	0.93	2.53E+09
26	Balolo-3.14-100	69	0.41	1.50E-02	0.59	2.17E+09
27	Buago-3-100+TFA	69	0.64	2.38E-02	0.36	1.09E+09
28	Eden-6-25	68	0.13	4.00E-03	0.87	9.67E+07
29	Tavaniko-5-50	68	0.68	2.50E-02	0.32	2.21E+09
30	Tavaniko-7-100	67	0.28	1.00E-02	0.72	1.10E+08
31	Buago-7-50	67	0.67	2.44E-02	0.33	2.07E+09
32	Tavarua-8-25	66	0.66	5.50E-02	0.33	1.19E+09
33	Eden-5-50	64	0.64	2.20E-02	0.36	2.55E+09
34	Tavarua-2-25	63	0.56	1.89E-02	0.44	1.22E+09
35	Tavaniko-7-100+TFA	62	0.27	9.00E-03	0.73	1.17E+08
36	Eden-5-100+TFA	62	0.47	1.50E-02	0.53	7.93E+08
37	Pinacple-5-25	58	0.52	1.60E-03	0.43	1.21E+09
38	Eden-5-25	57	0.63	1.90E-02	0.37	1.10E+09
39	Buago-3-50	56	0.59	1.79E-02	0.41	1.17E+09
40	Balolo-3.14-50	53	0.51	1.46E-02	0.49	2.95E+09
41	Tavaniko-7-50	48	0.65	1.63E-02	0.35	9.03E+08
42	Tavaniko-4-50	47	0.67	1.57E-02	0.38	6.86E+08
43	Buago-7-25	41	0.66	1.46E-02	0.34	7.55E+07
44	Eden-4-50	36	0.64	1.21E-02	0.36	7.44E+08
45	Balolo-3.14-25	35	0.40	7.50E-03	0.60	1.62E+09
46	Eden-4-100+TFA	33	0.42	7.36E-03	0.58	5.41E+08
47	Buago-3-25	30	0.27	4.33E-03	0.73	5.82E+08
48	Namotu-2-50	30	0.60	9.00E-03	0.40	5.49E+09
49	Eden-5-100	30	0.63	1.00E-02	0.37	6.70E+08
50	Tavaniko-5-25	25	0.28	4.00E-03	0.72	3.50E+08
51	Eden-4-100	25	0.48	6.30E-03	0.52	5.15E+07
52	Pinacple-5-100+TFA	23	0.26	3.00E-04	0.74	1.67E+08
53	Tavaniko-7-25	23	0.39	4.00E-03	0.61	1.67E+07
54	Tavaniko-4-25	20	0.35	3.79E-03	0.65	5.43E+08
55	Eden-4-25	18	0.56	5.30E-03	0.44	1.34E+08
56	Namotu- 2-25	15	0.33	2.60E-03	0.67	1.14E+08

Table S36. Complexity metric data for all fractions arranged from highest to lowest.

No	Sample code	Sample complexity	Chemical novelty	Sample Media contamination	Sample novelty	Sample diversity
1	Buago-7-100	126	0.39	2.65E-02	0.61	9.76E+09
2	Pinacple-5-100	140	0.33	2.49E-02	0.67	9.54E+09
3	Lau-1-50	79	0.76	3.20E-02	0.24	9.08E+09
4	Pinacple-5-50	96	0.61	3.20E-02	0.39	7.97E+09
5	Tavarua-2-100	161	0.63	5.50E-02	0.37	6.90E+09
6	Tavarua-2-100+TFA	153	0.61	5.10E-02	0.39	6.15E+09
7	Namotu-2-50	30	0.60	9.00E-03	0.40	5.49E+09
8	Eden-6-100	135	0.18	1.26E-02	0.82	3.39E+09
9	Eden-6-50	109	0.11	6.00E-03	0.89	3.31E+09
10	Tavarua-8-100+TFA	111	0.49	2.93E-02	0.51	3.23E+09
11	Buago-7-100+TFA	113	0.52	3.19E-02	0.48	3.13E+09
12	Balolo-3.14-50	53	0.51	1.46E-02	0.49	2.95E+09
13	Tavarua-8-50	108	0.66	3.73E-02	0.34	2.71E+09
14	Eden-5-50	64	0.64	2.20E-02	0.36	2.55E+09
15	Tavarua-8-100	69	0.07	6.00E-02	0.93	2.53E+09
16	Tavarua-2-50	74	0.66	3.30E-02	0.15	2.52E+09
17	Namotu-2-100	70	0.27	1.00E-02	0.73	2.26E+09
18	Tavaniko-5-50	68	0.68	2.50E-02	0.32	2.21E+09
19	Balolo-3.14-100	69	0.41	1.50E-02	0.59	2.17E+09
20	Buago-7-50	67	0.67	2.44E-02	0.33	2.07E+09
21	Eden-6-100+TFA	108	0.52	3.00E-02	0.47	2.06E+09
22	Tavaniko-5-100	94	0.37	1.89E-02	0.63	2.06E+09
23	Balolo-3.14-25	35	0.40	7.50E-03	0.60	1.62E+09
24	Namotu-2-100+TFA	70	0.24	9.00E-03	0.76	1.47E+09
25	Lau-1-25	70	0.24	9.00E-03	0.76	1.47E+09
26	Tavarua-2-25	63	0.56	1.89E-02	0.44	1.22E+09
27	Pinacple-5-25	58	0.52	1.60E-03	0.43	1.21E+09
28	Tavarua-8-25	66	0.66	5.50E-02	0.33	1.19E+09
29	Buago-3-50	56	0.59	1.79E-02	0.41	1.17E+09
30	Eden-5-25	57	0.63	1.90E-02	0.37	1.10E+09
31	Buago-3-100+TFA	69	0.64	2.38E-02	0.36	1.09E+09
32	Tavaniko-4-100+TFA	75	0.24	9.74E-03	0.76	9.44E+08
33	Lau-1-100	79	0.72	3.08E-02	0.28	9.08E+08
34	Tavaniko-7-50	48	0.65	1.63E-02	0.35	9.03E+08
35	Lau-1-100+TFA	96	0.54	4.00E-02	0.46	8.14E+08
36	Eden-5-100+TFA	62	0.47	1.50E-02	0.53	7.93E+08
37	Eden-4-50	36	0.64	1.21E-02	0.36	7.44E+08
38	Tavaniko-4-50	47	0.67	1.57E-02	0.38	6.86E+08
39	Eden-5-100	30	0.63	1.00E-02	0.37	6.70E+08
40	Tavaniko-4-100	95	0.33	1.68E-02	0.67	5.97E+08
41	Buago-3-25	30	0.27	4.33E-03	0.73	5.82E+08
42	Tavaniko-4-25	20	0.35	3.79E-03	0.65	5.43E+08
43	Eden-4-100+TFA	33	0.42	7.36E-03	0.58	5.41E+08
44	Tavaniko-5-25	25	0.28	4.00E-03	0.72	3.50E+08
45	Balolo-3.14-100+TFA	92	0.17	8.66E-02	0.83	2.70E+08
46	Tavaniko-5-100+TFA	94	0.37	1.89E-02	0.63	2.06E+08
47	Pinacple-5-100+TFA	23	0.26	3.00E-04	0.74	1.67E+08
48	Buago-3-100	88	0.56	2.65E-02	0.44	1.38E+08
49	Eden-4-25	18	0.56	5.30E-03	0.44	1.34E+08
50	Tavaniko-7-100+TFA	62	0.27	9.00E-03	0.73	1.17E+08
51	Namotu- 2-25	15	0.33	2.60E-03	0.67	1.14E+08
52	Tavaniko-7-100	67	0.28	1.00E-02	0.72	1.10E+08
53	Eden-6-25	68	0.13	4.00E-03	0.87	9.67E+07
54	Buago-7-25	41	0.66	1.46E-02	0.34	7.55E+07
55	Eden-4-100	25	0.48	6.30E-03	0.52	5.15E+07
56	Tavaniko-7-25	23	0.39	4.00E-03	0.61	1.67E+07

Table S37. Diversity metric data for all samples arranged from highest to lowest.

No	Date	Number	Genus	Site	Coordinates	Depth/m
1	23-03-2009	Pinnacles-5	Tunicate	Pinnacles	S17.48.041 E177.11.914	6
2	19-03-2009	Balolo-3.14	Didemnidae	Balolo	S17.10.750 E177.10.302	12
3	17-03-2009	Tavarua-8.9	Eudistoma	Tavarua	S17.51.752 E177.11.525	6-12
4	19-03-2009	Buago-7	Fascaplysinsopsis	Buago	S17.13.385 E177.12.032	4
5	19-03-2009	Buago-3	Porifera	Buago	S17.13.385 E177.12.032	11
6	17-03-2009	Tavuniko-4	Didemnidae	Tavuniko	S17.12.754 E177.10.219	8
7	17-03-2009	Tavuniko-5	Synoicum	Tavuniko	S17.12.754 E177.10.219	8
8	17-03-2009	Tavuniko-7	Liosina	Tavuniko	S17.12.754 E177.10.219	8
9	20-03-2009	Eden-5	Liosina	Eden	S17.11.824 E177.10.114	6-12
10	20-03-2009	Eden-4	Dysidea	Eden	S17.11.824 E177.10.114	10
11	20-03-2009	Eden-6	Pseudoceratina	Eden	S17.11.824 E177.10.114	18
12	24-03-2009	Namotu-2	Haplosclerida	Namotu	S17.50.682 E177.10.518	5
13	24-06-2009	Lau-1	Lissoclinum	Lau	not measured	reef
14	24-03-2009	Tavarua-2	Tunicate	Tavarua	S17.51.752 E177.11.525	20

Table S38. Fijian marine sponges and ascidians used in this study.

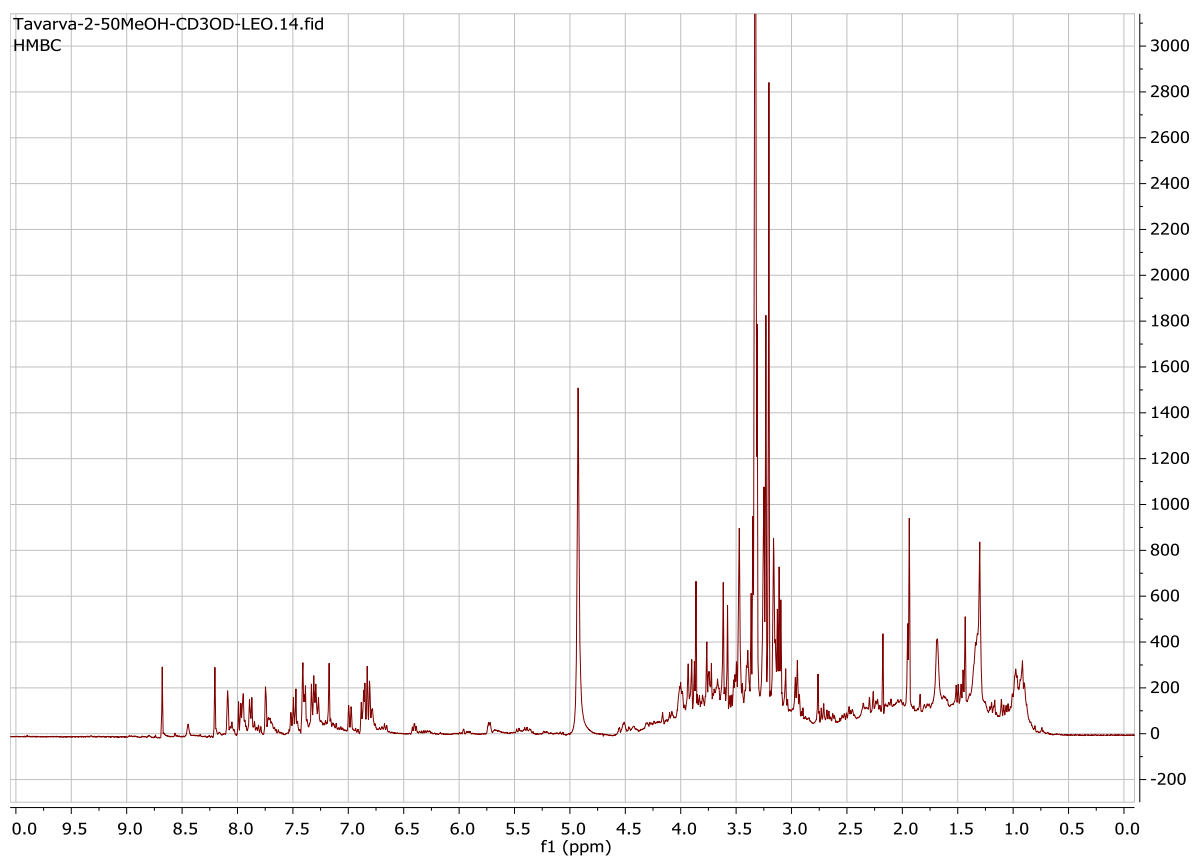


Figure S39. ^1H NMR (600 MHz, CD_3OD) of tavarua-2-SPE-50.

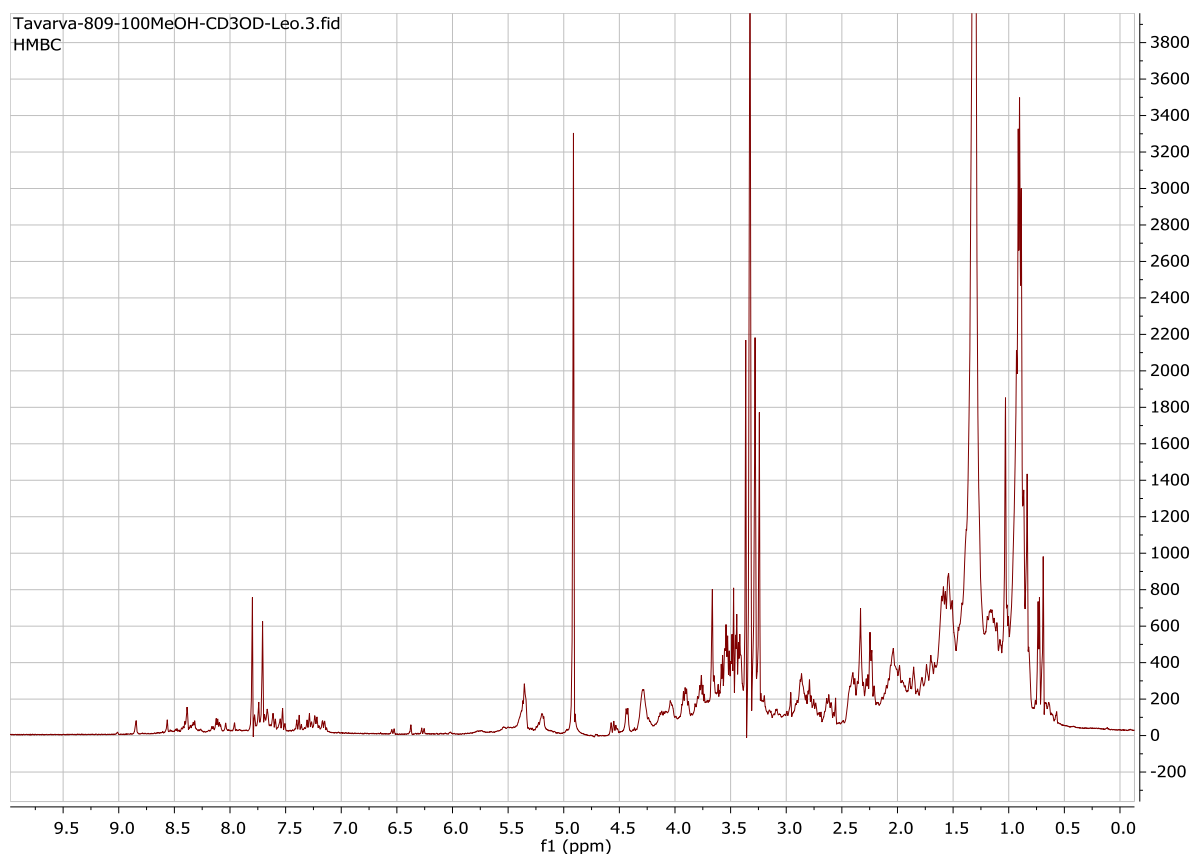


Figure S40. ^1H NMR (600 MHz, CD_3OD) of tavarua-8-SPE-100.

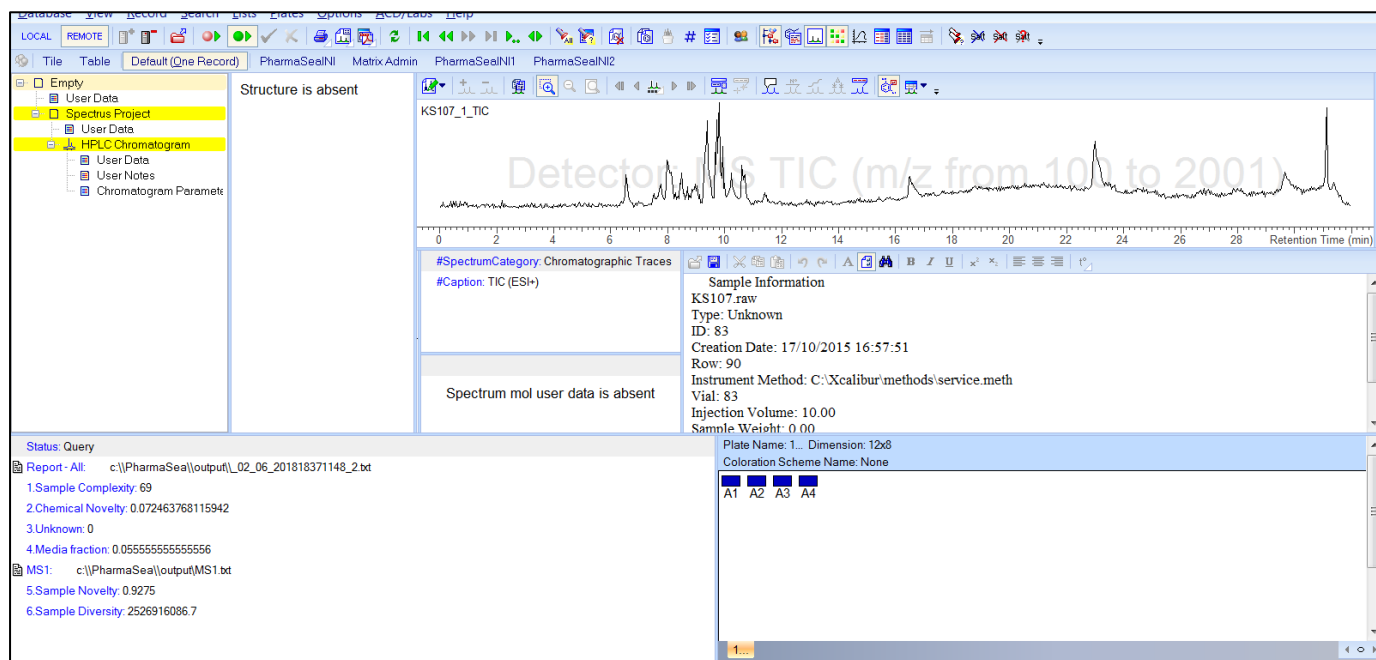


Figure S41. Metric data for tavarua-8-100.

index	Notation	Adjusted PTM	PTM	RT	Area (counts)	Height (counts)
1	[M+H] ⁺ U ⁺ u	351.075531	352.083351	6.523425	1660144	376932.5938
2	[M+H] ⁺ U ⁺ u	245.979178	246.986998	6.558092	21958654	5713015
3	[M+H] ⁺ U ⁺ u	199.050199	200.058019	6.558092	6383497	1115537.25
4	[M+H] ⁺ U ⁺ u	167.059789	168.067609	6.592858	27557340	4629797
5	[M+H] ⁺ U ⁺ u	458.994566	460.002386	7.034093	5870604	1041386.938
6	[M+H] ⁺ U ⁺ u	819.375596	820.383416	7.476593	1506554.875	300074.5938
7	[M+H] ⁺ U ⁺ u	335.080747	336.088567	7.580598	7679499	1052644.625
8	[M+H] ⁺ U ⁺ u	543.974935	544.982755	7.682425	2181033.75	435865.5938
9	[M+H] ⁺ U ⁺ u	258.960671	259.968491	7.716265	5690474.5	1187900.75
10	[M+H] ⁺ U ⁺ u	339.90278	340.9106	7.784423	7358570	1550246.5
11	[M+H] ⁺ U ⁺ u	430.983397	431.991217	7.955592	2876905.75	665570.625
12	[M+H] ⁺ U ⁺ u	333.065195	334.073015	8.024932	10672970	3508408.5
13	[M+H] ⁺ U ⁺ u	262.009548	263.017368	8.0921	30204774	4817765.5
14	[M+H] ⁺ U ⁺ u	333.065195	334.073015	8.161424	17182652	3563016.25
15	[M+H] ⁺ U ⁺ u	298.89316	299.90098	8.330928	4864815.5	969757.0625
16	[M+H] ⁺ U ⁺ u	610.975036	611.982856	8.365592	719893.5625	134619.2656
17	[M+H] ⁺ U ⁺ u	345.082112	346.089932	8.502091	33488144	8106570
18	[M+H] ⁺ U ⁺ u	262.009548	263.017368	8.53559	4729834.5	722122.3125
19	[M+H] ⁺ U ⁺ u	345.083182	346.091002	8.671092	10027896	2913569.5
20	[M+H] ⁺ U ⁺ u	575.041401	576.049221	9.008092	3595040.75	673526.75
21	[M+H] ⁺ U ⁺ u	440.987249	441.995069	9.247758	7013977	1130771.75
22	[M+H] ⁺ U ⁺ u	412.990307	413.998127	9.349767	54229064	11477296
23	[M+H] ⁺ U ⁺ u	432.096542	433.104362	9.417923	9001578	1776447.125
24	[2M+H] ⁺ U ⁺ u	424.9913755	850.990571	9.7551	5531047.5	934482.625
25	[M+H] ⁺ U ⁺ u	424.991665	425.999485	9.786925	208942464	28739034
26	[M+H] ⁺ U ⁺ u	404.945426	405.953246	9.959092	1547100.25	258452.4063
27	[M+H] ⁺ U ⁺ u	392.929524	393.937344	10.130425	5339217.5	1152776.5
28	[M+H] ⁺ U ⁺ u	406.955609	407.963429	10.266925	13391591	1879893
29	[M+H] ⁺ U ⁺ u	404.945426	405.953246	10.335092	3967325.5	696406.4375
30	[M+H] ⁺ U ⁺ u	427.007985	428.015805	10.677425	28482246	4828745
31	[M+H] ⁺ U ⁺ u	341.899063	342.906883	11.734591	2136498	400394.9375
32	[M+H] ⁺ U ⁺ u	452.949109	453.956929	11.968923	2597791.25	488049.2813
33	[M+H] ⁺ U ⁺ u	372.298343	373.306163	13.350759	2974327.75	563166.625
34	[M+H] ⁺ U ⁺ u	607.365691	608.373511	13.693093	1608328.5	216714.2813
35	[M+H] ⁺ U ⁺ u	317.292673	318.300493	13.893102	1566418.125	373956.0625
36	[M+H] ⁺ U ⁺ u	930.477327	931.485147	14.30159	1241937.875	166731.2813
37	[M+H] ⁺ U ⁺ u	633.381646	634.389466	14.36994	766139.0625	126682.6094
38	[M+H] ⁺ U ⁺ u	392.929524	393.937344	14.643925	1812874.375	262893.0625
39	[M+H] ⁺ U ⁺ u	621.381653	622.389473	14.710092	2870420	545269.125
40	[M+H] ⁺ U ⁺ u	707.450181	708.458001	15.419927	5765190.5	1085413.375
41	[M+H] ⁺ U ⁺ u	511.363472	512.371292	15.76009	10111053	1350695
42	[M+H] ⁺ U ⁺ u	483.292463	484.300283	16.278425	5680386.5	572279.5625
43	[2M+H] ⁺ U ⁺ u	481.353009	963.713838	16.547598	706984.125	146481.0781
44	[M+H] ⁺ U ⁺ u	481.349966	482.357786	16.582266	94614448	11065951

45	[M+Na] ⁺ U ⁺ u	481.351171	504.340941	16.582266	18587770	2112535
46	[M+H] ⁺ U ⁺ u	495.364777	496.372597	17.295933	7182372	732751.1875
47	[M+H] ⁺ U ⁺ u	497.308597	498.316417	17.432423	5326241	324508.3438
48	[M+H] ⁺ U ⁺ u	495.365125	496.372945	17.671089	12612388	1506559.25
49	[M+H] ⁺ U ⁺ u	507.365635	508.373455	17.806589	850076.4375	149813.0781
50	[M+H] ⁺ U ⁺ u	539.391472	540.399292	17.979256	1908884	258567.9531
51	[M+H] ⁺ U ⁺ u	551.392477	552.400297	18.767092	9358112	1309657.5
52	[M+H] ⁺ U ⁺ u	549.231335	550.239155	19.752258	642852.9375	144972.1406
53	[M+H] ⁺ U ⁺ u	537.270416	538.278236	20.296934	1077273.125	184238.7031
54	[M+H] ⁺ U ⁺ u	535.253483	536.261303	20.500925	1462297.875	281448.75
55	[M+H] ⁺ U ⁺ u	636.294562	637.302382	21.245926	1051355.25	166595.6875
56	[M+H] ⁺ U ⁺ u	580.320407	581.328227	21.721926	2828499.5	391619.3438
57	[M+H] ⁺ U ⁺ u	598.491887	599.499707	22.128597	1531759.625	230959.8594
58	[M+H] ⁺ U ⁺ u	522.316671	523.324491	23.039089	162357840	17677402
59	[M+H] ⁺ U ⁺ u	626.521472	627.529292	23.48176	3032503.75	395920.9063
60	[M+H] ⁺ U ⁺ u	652.53806	653.54588	23.616425	5761153.5	704911.1875
61	[M+Na] ⁺ U ⁺ u	652.538404	675.528174	23.616425	1254150.75	151980.875
62	[M+H] ⁺ U ⁺ u	523.317427	524.325247	24.153423	18118994	753079
63	[M+H] ⁺ U ⁺ u	928.215755	929.223575	24.324589	297247.9063	40591.86719
64	[M+H] ⁺ U ⁺ u	550.345003	551.352823	24.457422	1215762.25	161233.2031

Table S42. Unidentified masses in tavarua-8-100 (masses not found in PharmaSeaDB. Mass for compound **1** (green).

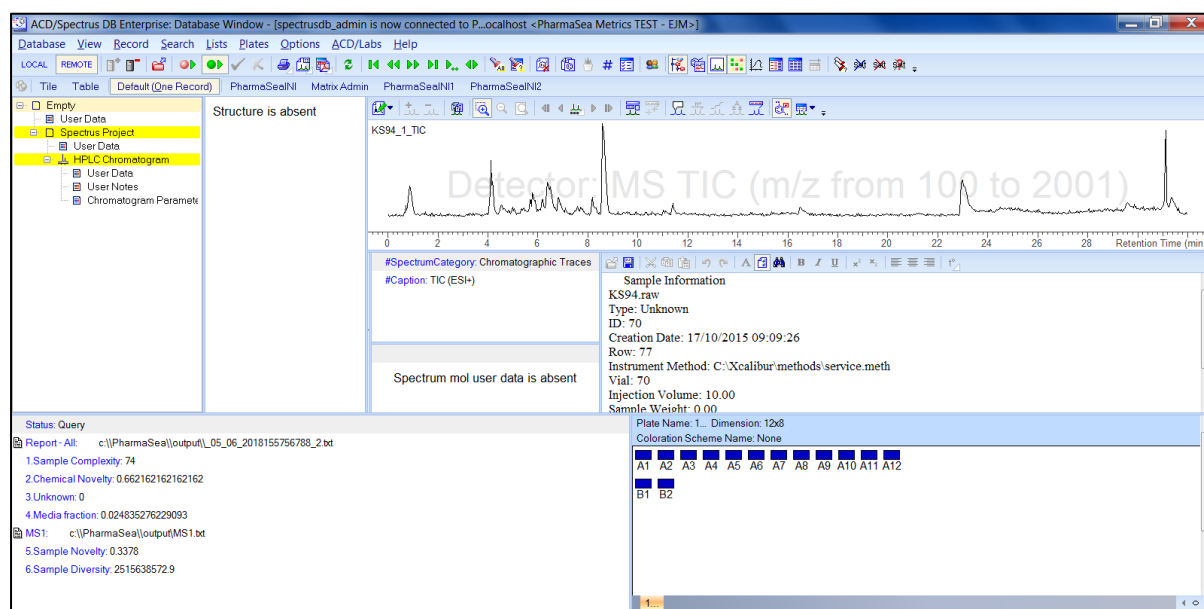


Figure S43. Metric data for tavarua-2-50 after processing with a mass range of 0.2 ppm.

index	Notation	Adjusted PTM	PTM	RT	Area (counts)	Height (counts)
1	[M+H] ⁺ U ⁺ u	289.142969	290.150789	3.92395	1863052.625	476671.125
2	[2M+H] ⁺ U ⁺ u	317.1469655	635.301751	4.149117	1833513.625	440470.0625
3	[M+H] ⁺ U ⁺ u	484.089259	485.097079	4.716784	2238138.5	483046.4375
4	[M+H] ⁺ U ⁺ u	489.084319	490.092139	4.75029	1419154.625	299124.75
5	[M+H] ⁺ U ⁺ u	581.200219	582.208039	5.017783	1961837.625	203842.9688
6	[M+H] ⁺ U ⁺ u	747.203381	748.211201	5.321126	531048.75	145307.5156
7	[2M+H] ⁺ U ⁺ u	377.1360635	755.279947	6.422616	2107612.75	395100.7813
8	[M+Na] ⁺ U ⁺ u	377.13637	400.12614	6.456125	16355297	1868236.375
9	[M+H] ⁺ U ⁺ u	468.269688	469.277508	7.57695	1095684.125	225208.7188
10	[2M+H] ⁺ U ⁺ u	507.1254485	1015.258717	8.190117	497354.8125	131687.5
11	[M+Na] ⁺ U ⁺ u	507.12734	530.11711	8.222616	11678545	2099523.75
12	[2M+H] ⁺ U ⁺ u	521.1429075	1043.293635	8.625283	10935016	2705634.25
13	[M+H] ⁺ U ⁺ u	485.254819	486.262639	8.758117	1677232.375	383905.0625
14	[M+H] ⁺ U ⁺ u	513.287319	514.295139	10.304461	2858998.25	551920.1875
15	[M+H] ⁺ U ⁺ u	606.256328	607.264148	10.872622	405220.4375	89519.46094
16	[M+H] ⁺ U ⁺ u	509.280301	510.288121	11.852622	2841848	507860.5625
17	[M+H] ⁺ U ⁺ u	504.215547	505.223367	15.565784	7794196.5	582190.125
18	[M+Na] ⁺ U ⁺ u	504.213262	527.203032	16.06695	9497951	582790.5
19	[M+Na] ⁺ U ⁺ u	481.351171	504.340941	16.573126	7843521.5	938882.8125
20	[M+H] ⁺ U ⁺ u	481.353006	482.360826	16.573126	37473716	4271565.5
21	[M+H] ⁺ U ⁺ u	682.19892	683.20674	20.509951	1118910.875	121562.4375
22	[M+H] ⁺ U ⁺ u	636.294562	637.302382	21.252451	1111424.75	174955.8594
23	[M+H] ⁺ U ⁺ u	670.500297	671.508117	22.745285	2481364.75	280006.6563
24	[M+H] ⁺ U ⁺ u	522.316671	523.324491	23.077616	207465504	20861098
25	[M+H] ⁺ U ⁺ u	670.500297	671.508117	24.899786	284310.1563	56117.85547

Table S44. New masses in tavarua-2-50 (masses not found in PharmaSeaDB) using a 0.2 ppm mass error at m/z 500. Masses of compounds **4** and **5** are highlighted.

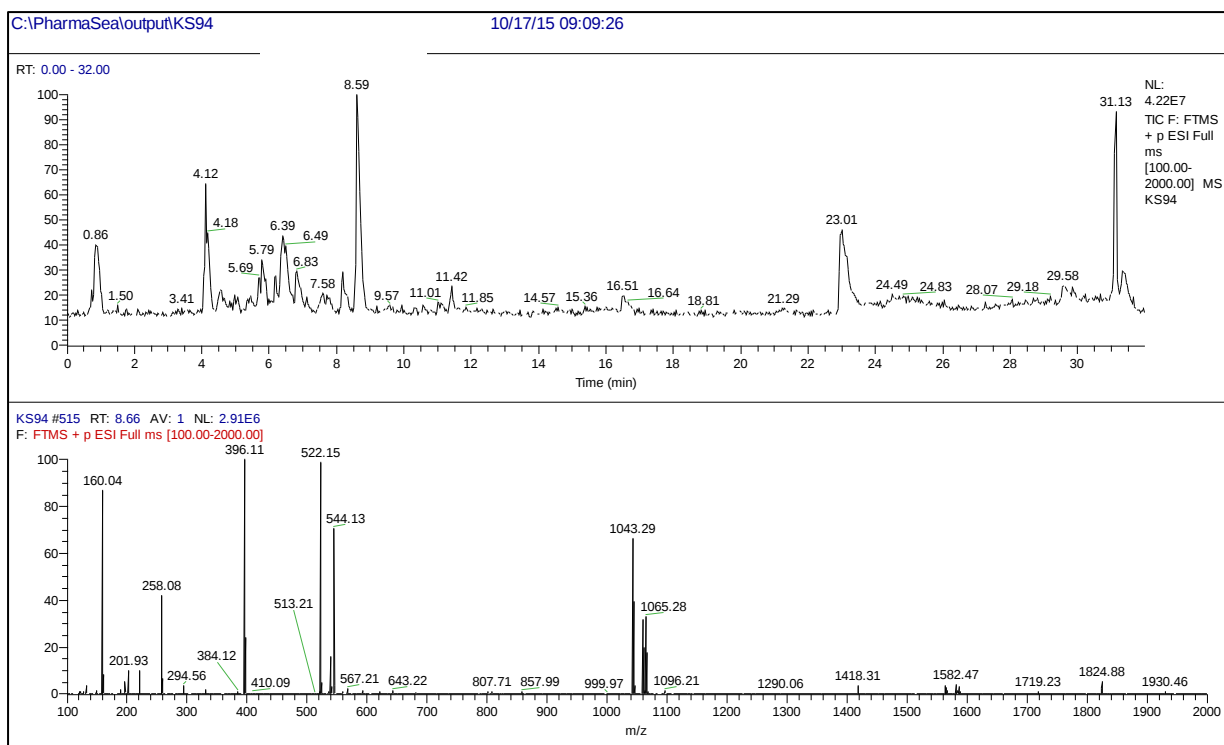


Figure S45. Total ion chromatogram (TIC) for tavarua-2-50 showing the MS spectrum of the peak at 8.59 minutes ($m/z = 522$) corresponding to compound **4**.

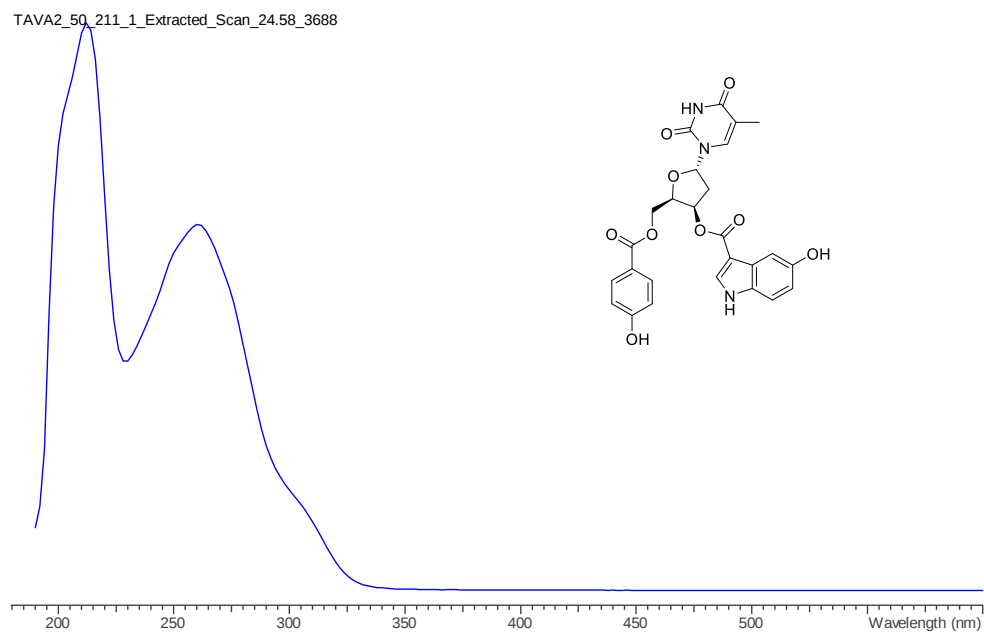


Figure S46. UV spectrum for compound **4**.

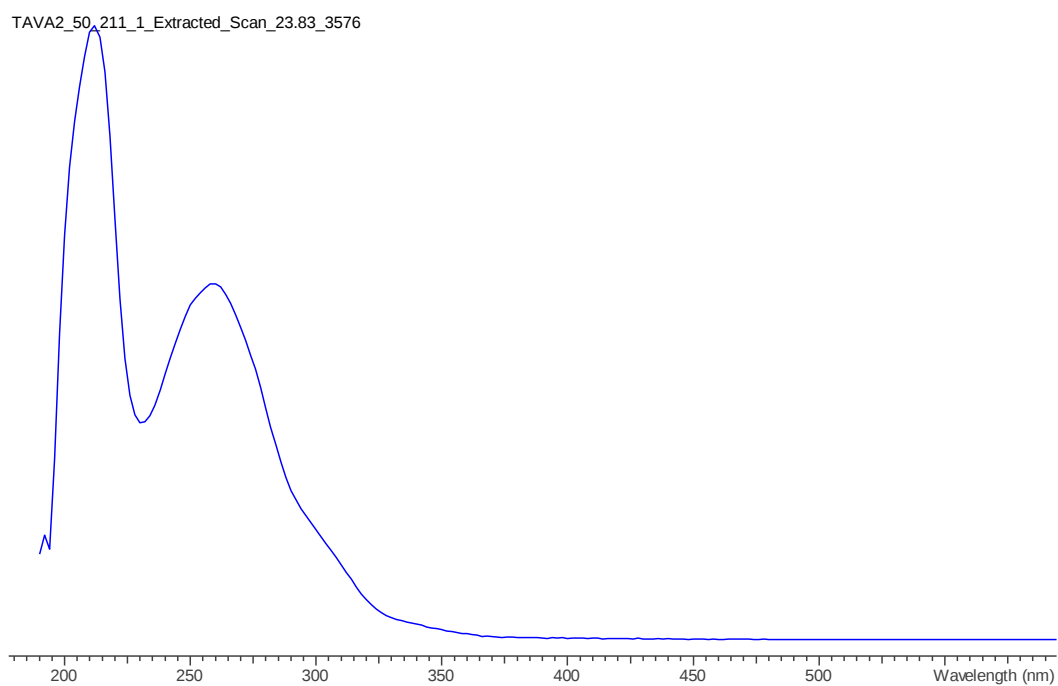


Figure S47. UV spectrum for compound **5**.

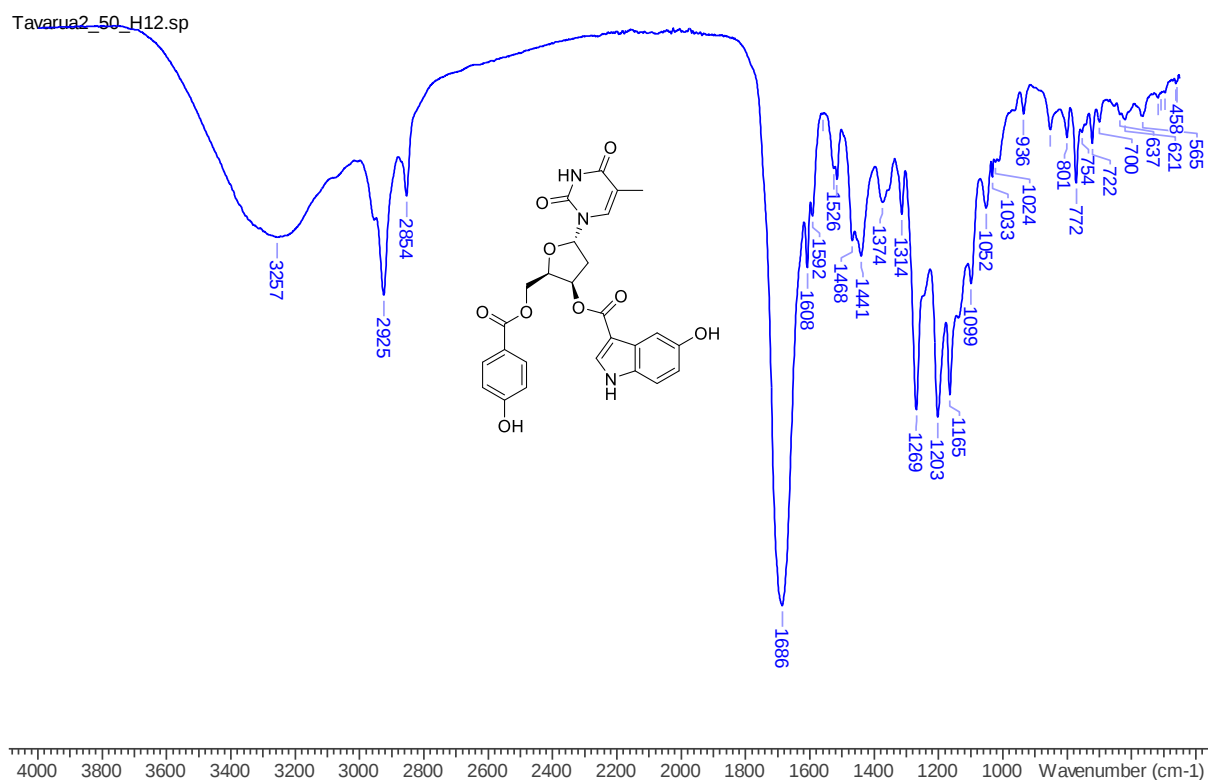


Figure S48. IR spectrum for compound **4**.

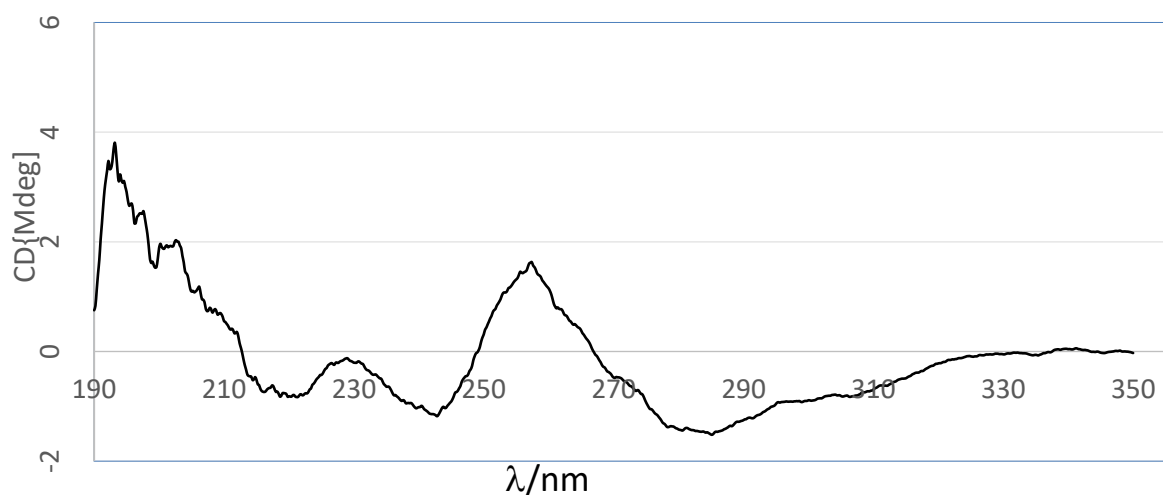


Figure S49. ECD spectrum for compound **4**. The spectrum represented an average of 3 scans collected at a scan speed of 10 nm/min, response of 2 seconds and bandwidth of 1 nm.

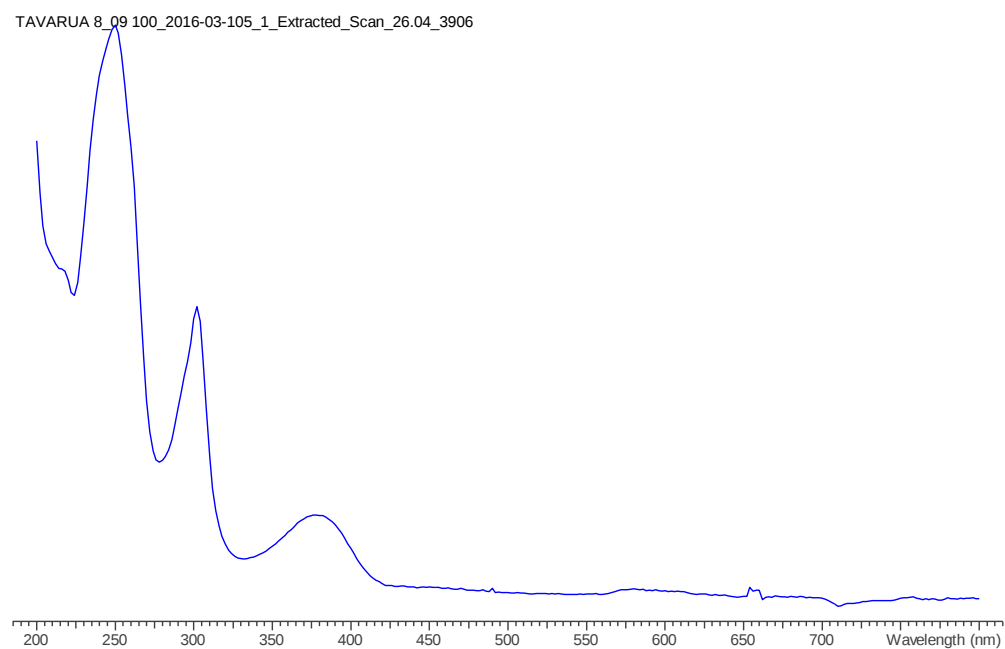


Figure S50. UV spectrum for compound **1**

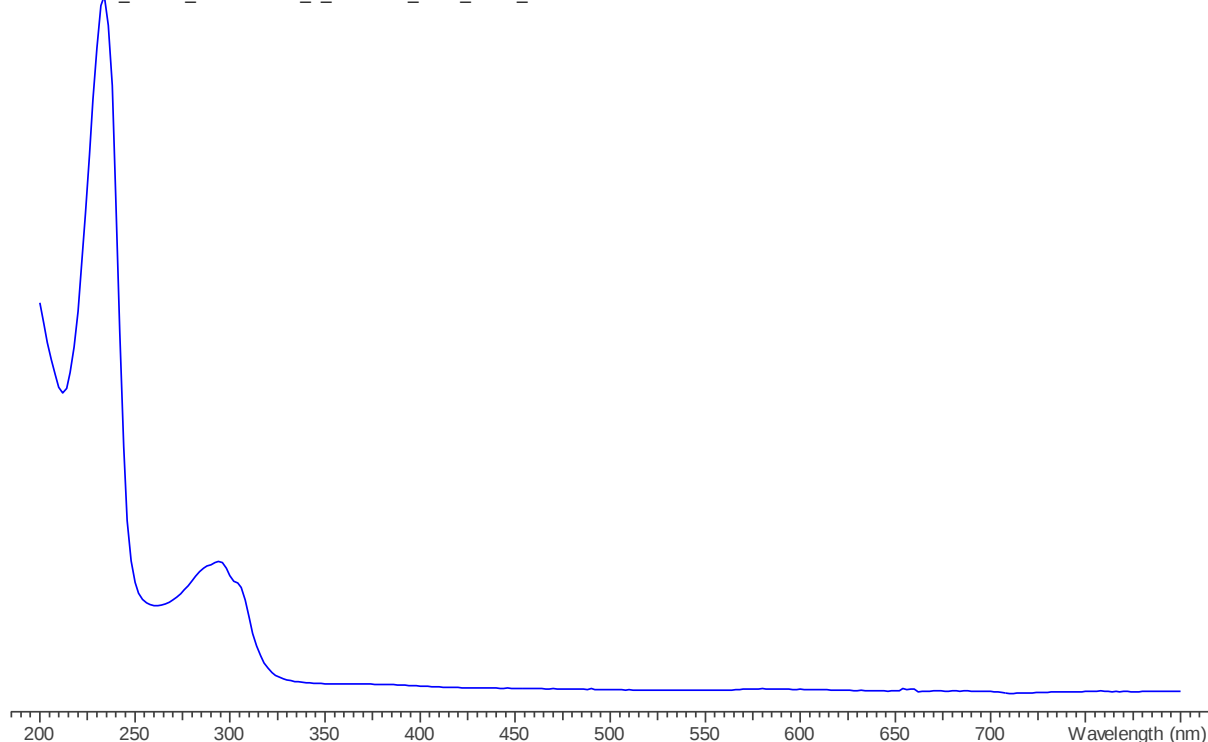


Figure S51. UV profile for compound **2**

Part B. Computational methods

Initial structures for the various stereoisomers of **4** were created using the Avogadro molecular modelling software version 1.1.1 and energy minimized with the MMFF94 force field.^{1,2} Compound **5** was not considered as the addition of a single methyl group was unlikely to result in any significant alteration to the ECD spectrum. The initial MMFF94 structures were then submitted to conformational searches using the Avogadro/OpenBabel interface.³ The weighted rotor search method was used and 500 conformers were generated for each structure again using the MMFF94 force field for energy minimization.

The lowest energy structures obtained from the conformational searches were then re-optimised at the DFT level of theory using the PBE exchange-correlation functional and the double- ζ Def2-SVP basis set.^{4,5} Solvation effects (methanol) were modelled with the conductor-like polarized continuum method (PCM).⁶ The resulting optimized structures were then used in TDDFT calculations of the ECD spectra of each species. The hybrid functional ω B97X was employed for the TDDFT calculations since range-separated approaches to the inclusion of exact Hartree-Fock exchange are broadly accepted as providing reliable reproduction of spectral energies and intensities.⁷ The Def2-SVPD basis set which includes diffuse functions and was designed to provide an improved description of optical response properties

was used and methanol solvation was again modelled using the PCM approach.⁸ All electronic structure calculations were performed using Orca version 4.0.0.⁹

The electronic excitations obtained from the TDDFT calculations were fit using Gaussian functions with a 0.1 eV broadening. The best match to the features of the experimental ECD spectrum was obtained for the *S,R,R* stereoisomer (C1',C3',C4') and the calculated spectrum intensity for this structure was normalized using the intensity of the experimental maximum at ~190 nm. This same scaling was applied to the remaining calculated spectra thus providing estimates of the peak intensities that would be expected in experimental ECD spectra of those species

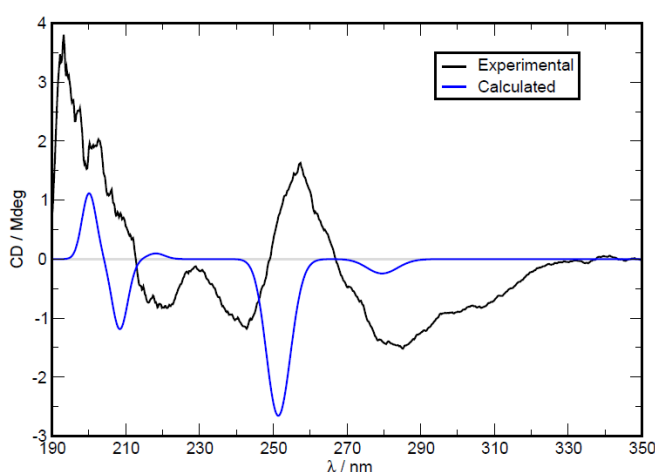


Figure S52. ECD spectrum overlay of experimental vs calculated (*R,R,R*) for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**).

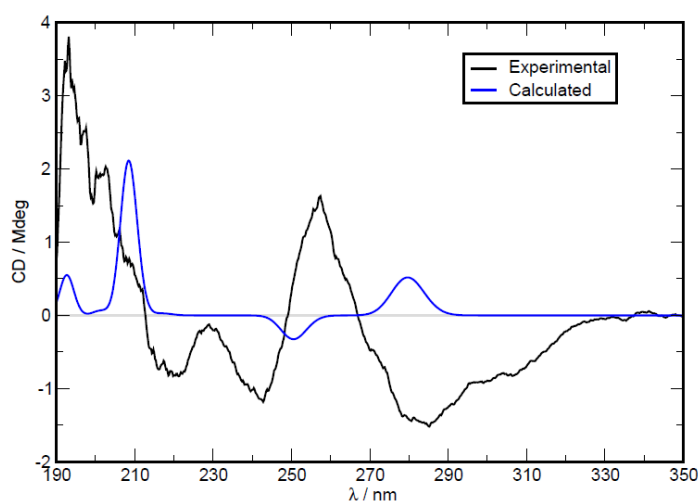


Figure S53. ECD spectrum overlay of experimental vs calculated (*R,R,S*) for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**).

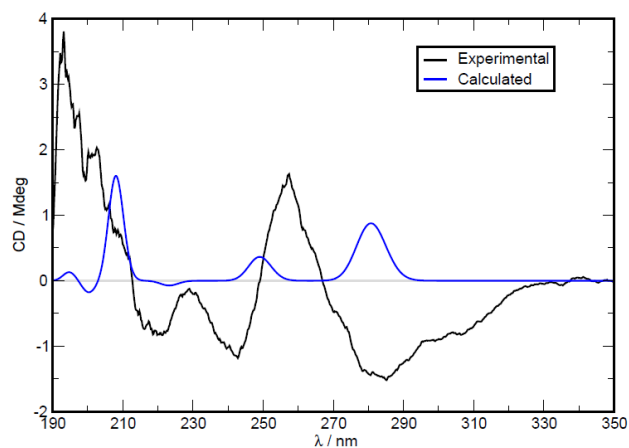


Figure S54. ECD spectrum overlay of experimental vs calculated (*R,S,R* for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**)).

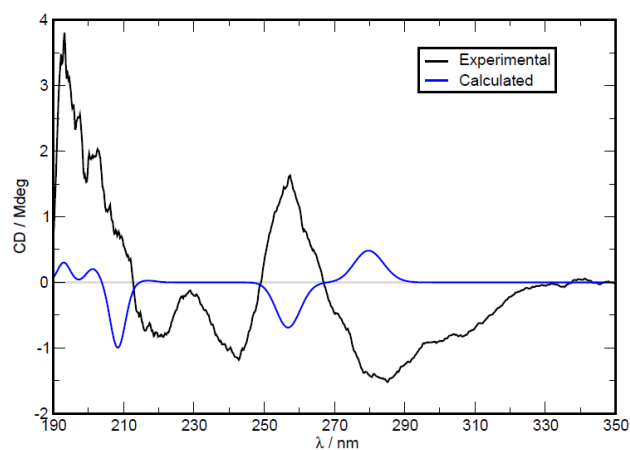


Figure S55. ECD spectrum overlay of experimental vs calculated (*R,S,S*) for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**)).

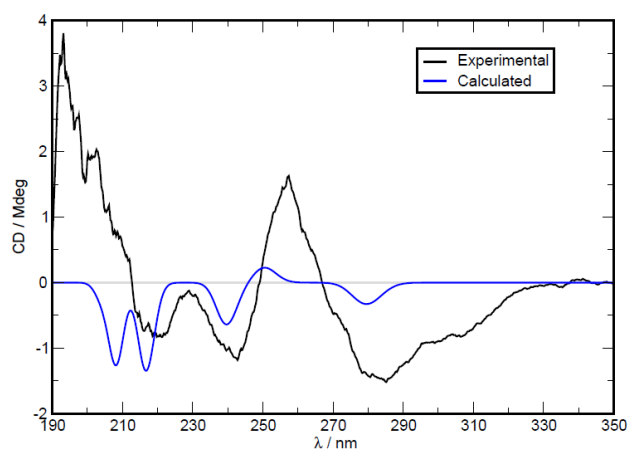


Figure S56. ECD spectrum overlay of experimental vs calculated (*S,R,S*) for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**)).

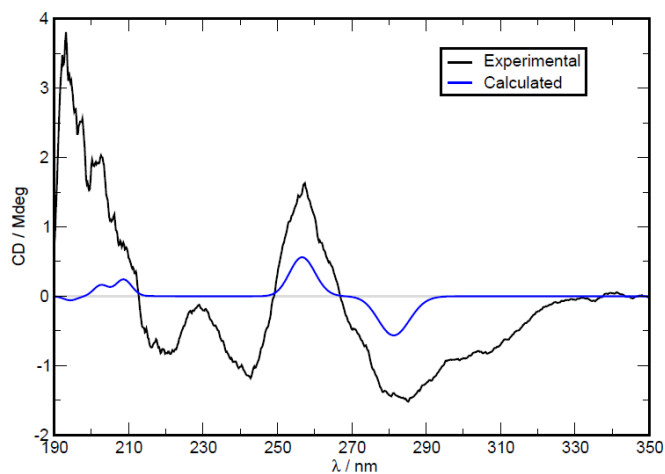


Figure S57. ECD spectrum overlay of experimental vs calculated (*S,S,R*) for C1',C3',C4' for tavarua deoxyriboside A (**4**).

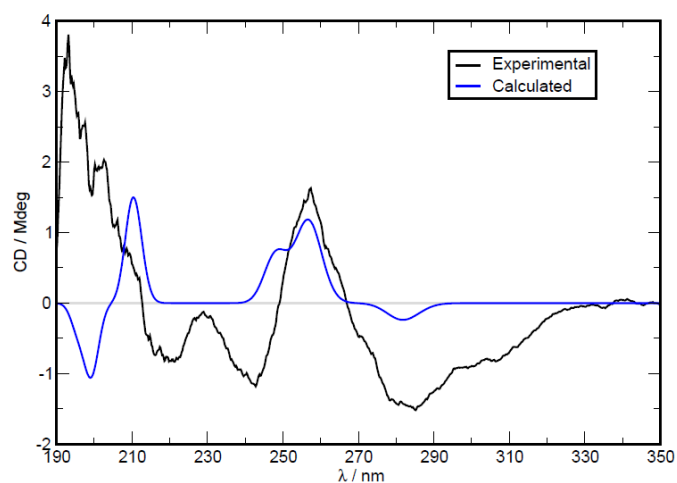


Figure S58. ECD spectrum overlay of experimental vs calculated (*S,S,S*) for C-1',C-3',C-4' for tavarua deoxyriboside A (**4**).

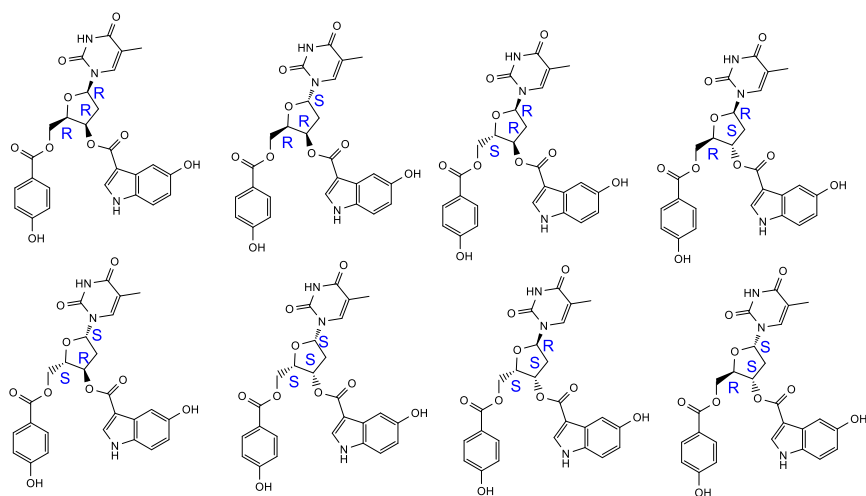


Figure S59. Structural configurations of the 8 possible stereoisomers for tavarua deoxyriboside A (**4**) determined by ACDChemsKetch.¹⁰

Part C. Ascidian Taxonomy

Molecular and Morphological Characterization of Two Ascidian Colonies (Tava2 and Tava8) from Fiji

Materials and Methods

The total DNA was extracted using a modified CTAB method¹¹ from about 10-15 zooids isolated from colonies preserved in RNA-later. A COI (Cytochrome Oxidase subunit I) gene fragment of about 850 bp, containing the 650 bp-long region commonly used as DNA barcode, was amplified using the primer pair dinF-Nux1R (dinF:5'-CGTTGRTTTATRTCTACWAATCATAARGA-3';Nux1R:5'-CAGTAAAATAWGCTCGRGARTC -3')¹². A 600 bp-long fragment of the 18S rRNA gene (18S rDNA) was amplified using the primer pair BF-AR (BF: 5'-CTGGTTGATCCTGCCAGTAGTBATAT-3';AR:5'-AGCTGGAATTACCGCGGCTGCT-3') designed as variants of already published ascidian 18S rRNA primers. In particular, AR is the reverse version of the 18SA-F primer¹³, while BF was designed starting from both the 18S-F1¹⁴ and 18S-F3¹⁵ primers.

Amplifications of the COI fragment were performed with the high fidelity PrimeStar HS DNA polymerase¹⁶ in a 25 µl reaction volume containing: 1X reaction buffer with 1 mM final concentration of MgCl₂¹⁶, 0.2 mM of each dNTP, 0.3 µM of each of the two primers and 1.25 Units of PrimeStar HS DNA polymerase.¹⁶ Following the manufacturer's protocol, the amplification conditions were: 30 cycles with denaturation for 10 s at 98°C, annealing for 15 s at 46°C, extension for 1 min 30 s at 72 °C; a final elongation step of 5 min at 72°C.

Amplifications of the 18S rDNA fragment were performed with the DreamTaq polymerase¹⁷ in a final volume of 25 µl containing: 1X reaction buffer with 2 mM final concentration of MgCl₂ (Thermo Scientific), 0.2 mM of each dNTP, 0.5 µM of each of the two primers, and 1.25 Units of DreamTaq polymerase.¹⁷ The amplification conditions were: an initial denaturation for 3 min at 95°C, then 30 amplification cycles (denaturation for 30 s at 95°C; annealing for 30 s at 60 °C; extension for 1 min 30 s at 72 °C) followed by a final elongation step of 5 min at 72°C.

Amplicons were purified using the DNA Clean & Concentrator kit¹⁸, and directly sequenced according to the Sanger method at the Eurofins Genomics (Ebersberg, Germany)¹⁹ or at the Microsynth AG (Switzerland)²⁰. Sequence quality check and assembly were carried out with Geneious ver. 5.5.7.^{21,21}

The obtained COI and 18S rDNA sequences were compared to the public NCBI nucleotide non-redundant database²² through the online Basic Local Alignment Search Tool (BlastN).^{23,24} Then, phylogenetic analyses within Aplousobranchia were carried out separately for COI and 18S rDNA, selecting from the nt-nr database only reliable homologous sequences having a length comparable to that of our samples, and including just 1-2 sequences for each species. Therefore, sequences too short or with clear erroneous taxonomic assignment were excluded from our final datasets, that include 131 aplousobranch sequences for COI and 48 for 18S rDNA.

COI sequences were aligned with MAFFT)²⁵, and the alignment was manually optimized preserving the codon structure. 18S rDNA sequences were aligned to the 18S secondary structure-based alignment of 110 deuterostomes published by Tsagkogeorga et al. (2009)¹⁴ and available on the Dryad Digital Repository²⁶. In particular, we used the MAFFT program²⁷ to add new aplousobranch sequences to a small version of the Tsagkogeorga's alignment consisting of only 26 taxa, that is two *Ciona* (Phlebobranchia) and 24 aplousobranch sequences. Therefore, our final 18S alignment is roughly based on the secondary structure. The final alignment of COI was 1602 nt long, while that of 18S was 846 nt long.

Phylogenetic reconstructions of Aplousobranchia were performed according to the maximum likelihood (ML) method using the online PHYML v3.0 software, which includes the automatic model selection by SMS (Smart Model Selection)^{28,29}. The best-fit substitution model selected under the Akaike Information Criterion (AIC) was the GTR+I+G for COI and TN93+G for the 18S. The proportion of invariant sites (I) and the gamma shape parameter (alpha) for the 4 rate categories were estimated by the PHYML v3.0 software itself. Bootstrap values, indicating node reliability, were based on 100 replicates. A total of 20 Phlebobranchia and Stolidobranchia species were used as outgroup in the COI analyses, while two *Ciona* species were set as outgroup in the 18S rDNA reconstructions (Table S64 and S65, respectively).

Morphological Analyses

Zooids of the Tava8 and Tava2 specimens were extracted from the tunic of colonies preserved in RNA-later. Unfortunately, due to the hard contraction of all samples, it was not possible to clearly see the inner characteristics of the zooids, therefore the specimen identification was possible only at level of family.

Tava 8 (Figure 60 a-d)

Phylum: Chordata

Subphylum: Tunicata

Class: Ascidiacea

Order: Enterogona

Suborder: Aplousobranchia

Family: Polycitoridae Michaelsen, 1904 or Pseudodistomidae Harant, 1931

Material examined. Tava8: small portion of a colony.

Description. The living small cushion-shaped colonies, up to 3 cm in height, appeared greenish in colour, with a thin whitish layer above and below externally to the test matrix (Figure 60 a-c). The

zooid body was divided in thorax and abdomen, up to 2 mm in length (Figure 60d), however the presence of a posterior abdomen was unclear. The very contracted zooids were embedded in the common test, not arranged in systems (Figure 60a). The oral and atrial apertures were lobed and opened at the level of the whitish external layers on both sides. The few clearly observed zooid anatomical features of the zooids are compatible with the assignment of the Tava8 specimen to the family Polycitoridae (zooids divided into thorax and abdomen) or to the family Pseudodistomidae (zooids divided into thorax, abdomen and posterior abdomen).

Tava2 (Figure 61 a-c)

Phylum: Chordata

Subphylum: Tunicata

Class: Ascidiacea

Order: Enterogona

Suborder: Aplousobranchia

Family: Didemnidae Giard, 1872

Material Examined. Tava2: small fragments of a colony.

Description. Encrusting colony constituted by a thin whitish and pink sheet with abundant star shaped aragonite spicules in the common tunic (Figure 61). Small zooids, up to 1.5 mm long, divided into thorax and abdomen. The pharynx was crossed by few rows of stigmata (Figure 61c). The oral siphon was lobed, while the atrial aperture appeared tubular and located near the posterior end of the thorax (Figure 61c).

Remarks. The family Didemnidae Giard, 1872 includes colonial ascidians characterized by the presence of star shape aragonite spicules in the common tunic. The zooids are very small (about 1 mm) and divided into thorax and abdomen. The pharynx has 3 or 4 rows of stigmata with the oral aperture lobed. Many genera belong to this family, such as *Trididemnum*, *Atrium* Kott, 1893, *Leptoclinides* Bjerkan, 1905, *Lissoclinum* Verrill, 1871, *Diplosoma* Macdonald, 1859, *Didemnum* Savigny 1816, *Clitella* Kott, 2001, *Polysyncraton* Nott, 1892. These genera are distinguished firstly for the number of rows of stigmata: three for the species belonging to the genus *Trididemnum*, and four for the other genera.³⁰ Based on the zooid body shape and the presence of spicules in the common tunic, it was possible to assign the Tava2 specimen only at family level, and exactly to the family Didemnidae.

Combining Molecular and Morphological Results

The Tava2 and Tava8 specimens were molecularly characterized using the mitochondrial COI gene and the nuclear 18S rDNA (Figure 63 and Figure 62, respectively). These genes were selected

since the COI has been demonstrated to be a reliable DNA barcode in metazoans³¹ and its sequence is available for many ascidian species, while the 18S rDNA has been used for reconstructing the ascidian phylogeny¹⁴ and its sequence is available for several aplousobranchs sampled in localities close to those of the Tava's origin.^{32,33}

Molecular data support the assignment of Tava8 to the genus *Eudistoma* (Polycitoridae) or *Pseudodistoma* (Pseudodistomidae). Indeed, in the COI phylogenetic tree of Aplousobranchia, Tava8 is positioned in a well-supported clade (bootstrap: 86%) including all 13 analysed *Eudistoma* and *Pseudodistoma* sequences. Unfortunately, the internal nodes of this clade remain unresolved, preventing further hypotheses on the assignment at genus or species level. In the 18S phylogenetic tree there are no *Pseudodistoma* sequences, so Tava8 clusters with the only two available *Eudistoma* species (bootstrap: 100%). It is important to notice that the close relationship between the genera *Eudistoma* and *Pseudodistoma* was already observed at both morphological³⁴ and molecular level.³⁵ Morphological analyses suggest the belonging of Tava8 to the family Polycitoridae or Pseudodistomidae, so we can conclude that molecular and morphological data point to the assignment of Tava8 to the genus *Eudistoma* (Polycitoridae) or *Pseudodistoma* (Pseudodistomidae).

Concerning Tava2, in the COI phylogenetic tree (Figure S63) this sequence is characterized by a long branch, so it clusters to fast-evolving sequences of the genus *Lissoclinum* (Didemnidae) but without statistical support (bootstrap: 33%). In the COI phylogenetic tree several genera are monophyletic and highly statistically supported (see Figure S63), but the basal nodes are mainly unresolved, so there are no clades corresponding to the different Aplousobranchia families (such as Didemnidae, Polycitoridae and Polyclinidae). Therefore, the COI does not even allow assigning Tava2 to a given family. On the contrary, in the 18S phylogenetic tree, Tava2 is well nested in the well-supported of the Didemnidae family (bootstrap: 75% and 91% with and without the genus *Leptoclinides*, respectively). In particular, Tava2 forms the basal node of a highly supported clade (bootstrap: 85%) including all but one *Didemnum* species (i.e., Ddsp.AB211072) plus only one of the five represented *Trididemnum* species (i.e., Trcy.KJ632950). So, unfortunately, we cannot unambiguously assign Tava8 to the genus *Didemnum*, although this is a reasonable hypothesis. In conclusion, the 18S rDNA data agree to the results of the morphological analyses, i.e., with the assignment of Tava2 to the family Didemnidae.

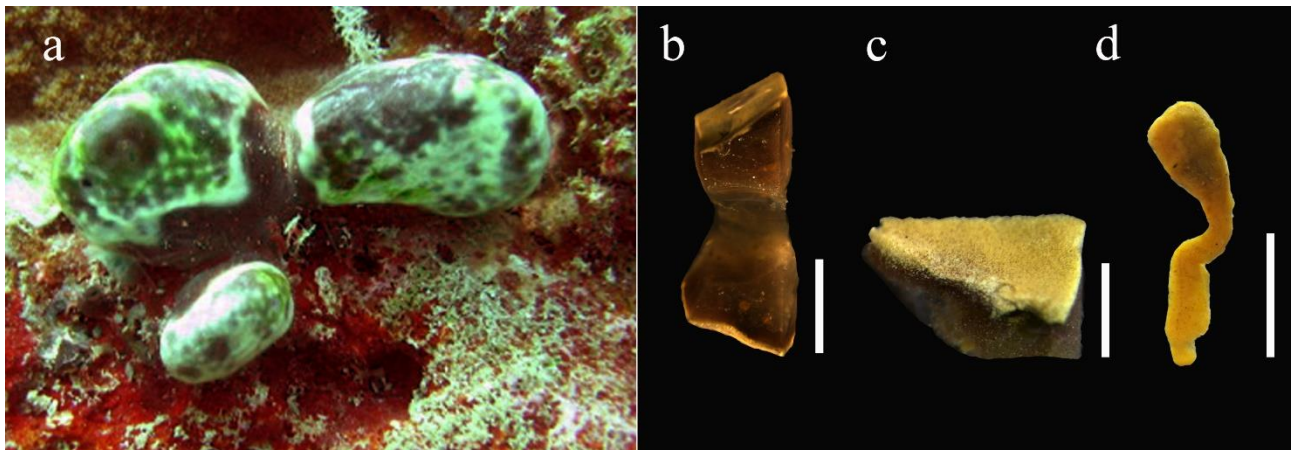


Figure S60. Tava8 specimen, Polycitoridae. (a) Live colony. (b) Preserved material showing the two whitish layers above and below the thick brownish matrix. (c) Detail of the whitish layer. (d) Zooid. Scale bars: 0.5 mm, 1 mm, 1 mm.

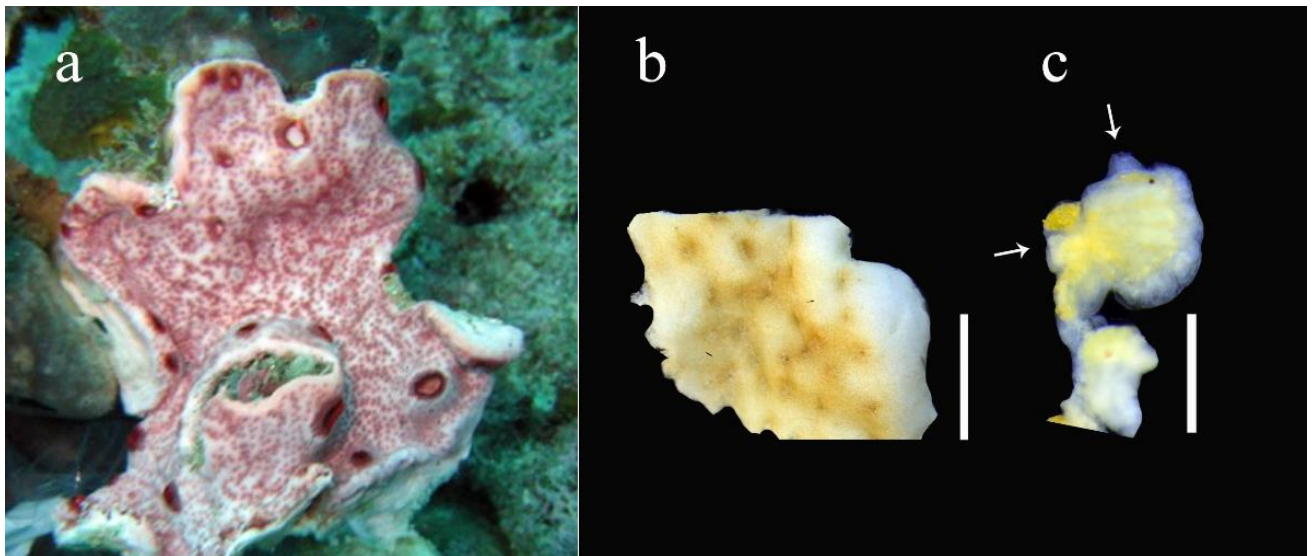


Figure S61. Tava2 specimen, Didemnidae (a) Live colony. (b) Preserved colony. (c) Zooids with white arrows pointing out the lobed oral siphon and the tubular atrial aperture. Scale bars: 2 mm, 0.5 mm.

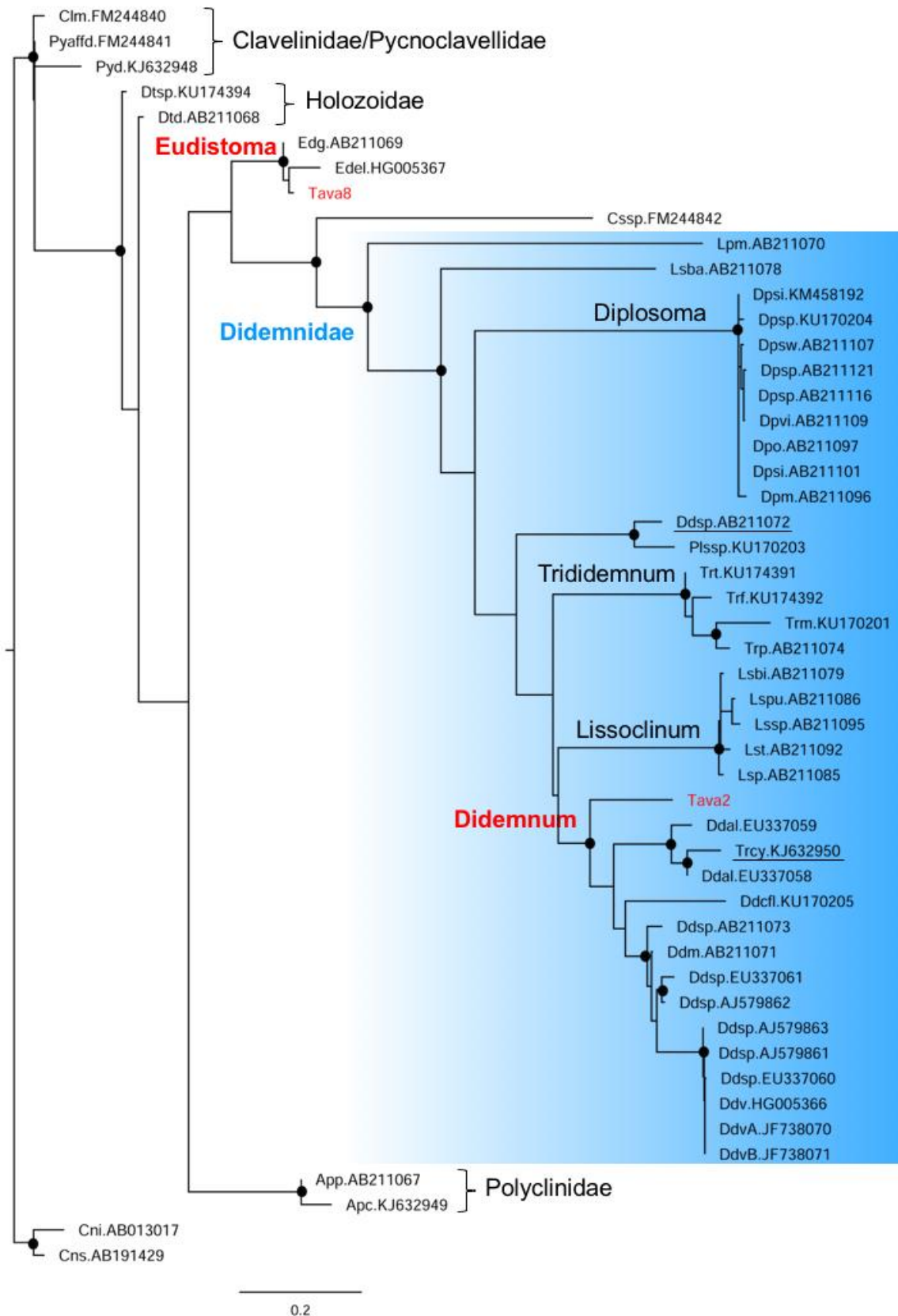


Figure S62. Maximum likelihood phylogenetic tree of Aplousobranchia based on a fragment of the nuclear 18S rDNA gene (PhyML: TN93 + G model; bootstrap on 100 replicates). Black dots indicate well supported nodes with bootstrap $\geq 75\%$. Two *Ciona* species were used as outgroups. Genus abbreviations are reported in Table S66. Analysed sequences are list in Table S65.

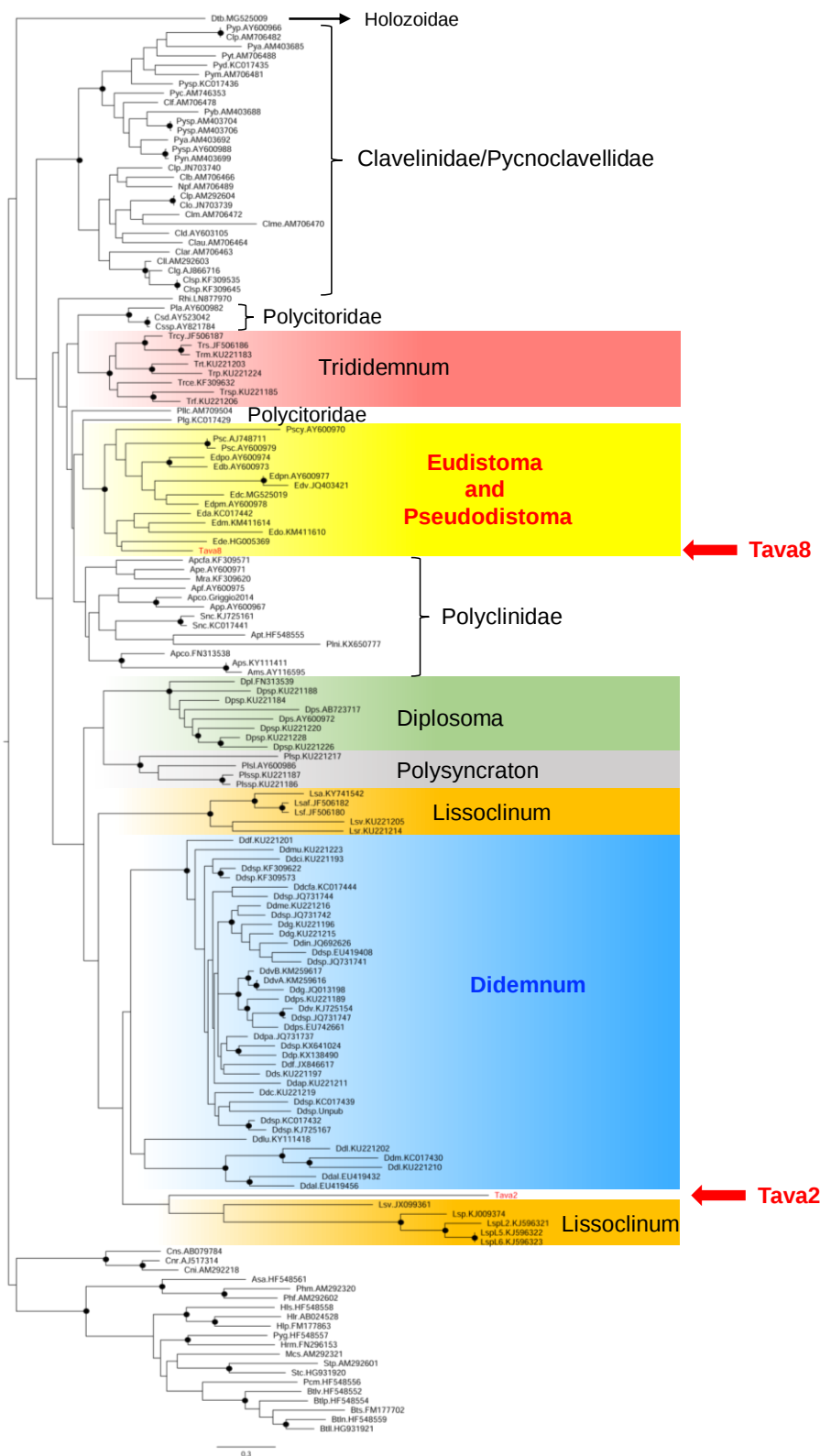


Figure S63. Maximum likelihood phylogenetic tree of Aplousobranchia based on a fragment of the mitochondrial COI gene (PhyML: GTR + I + G model; bootstrap on 100 replicates). Black dots indicate well supported nodes with bootstrap $\geq 75\%$. A total of 20 Phebobranchia and Stolidobranchia species were used as outgroups. Genus abbreviations are reported in Table S66. Analysed sequences are listed in Table S64.

Order	Family	Species	Description Line	bp	AC
Aplousobranchia	Clavelinidae	Clavelina arafurensis	Clavelina arafurensis coi gene for cytochrome c oxid	580	AM706463
Aplousobranchia	Clavelinidae	Clavelina australis	Clavelina australis coi gene for cytochrome c oxidase	635	AM706464
Aplousobranchia	Clavelinidae	Clavelina breve	Clavelina breve coi gene for cytochrome c oxidase,h	571	AM706466
Aplousobranchia	Clavelinidae	Clavelina delavallei	Clavelina delavallei cytochrome c oxidase subunit I	617	AY603105
Aplousobranchia	Clavelinidae	Clavelina flava	Clavelina flava coi gene for cytochrome c oxidase, h	602	AM706478
Aplousobranchia	Clavelinidae	Clavelina gemmae	Clavelina gemmae mitochondrial partial coi gene for	600	AJ866716
Aplousobranchia	Clavelinidae	Clavelina lepadiformis	entire mtDNA and COI		AM292603
Aplousobranchia	Clavelinidae	Clavelina meridionalis	Clavelina meridionalis coi gene for cytochrome c oxid	637	AM706470
Aplousobranchia	Clavelinidae	Clavelina moluccensis	Clavelina moluccensis coi gene for cytochrome c oxid	685	AM706472
Aplousobranchia	Clavelinidae	Clavelina oblonga	Clavelina oblonga haplotype 2 cytochrome c oxidase	672	JN703739
Aplousobranchia	Clavelinidae	Clavelina phlegraea	entire mtDNA and COI		AM292604
Aplousobranchia	Clavelinidae	Clavelina picta	Clavelina picta cytochrome c oxidase subunit I gene	629	JN703740
Aplousobranchia	Clavelinidae	Clavelina producta	Clavelina producta coi gene for cytochrome c oxidase	537	AM706482
Aplousobranchia	Clavelinidae	Clavelina sp. AM-4	Clavelina sp. AM-4 cytochrome oxidase subunit I (C	586	KF309535
Aplousobranchia	Clavelinidae	Clavelina sp. SCR-5	Clavelina sp. SCR-5 cytochrome oxidase subunit I (C	586	KF309645
Aplousobranchia	Diazonidae	Rhopalaea idoneta	entire mtDNA and COI		LN877970
Aplousobranchia	Didemnidae	Didemnum albidum	Didemnum albidum voucher RMNH UROCH 750 cyto	586	EU419432
Aplousobranchia	Didemnidae	Didemnum albidum	Didemnum albidum voucher RMNH UROCH 751 cyto	586	EU419456
Aplousobranchia	Didemnidae	Didemnum apurto	Didemnum apurto voucher UF 813 cytochrome oxid	604	KU221211
Aplousobranchia	Didemnidae	Didemnum cf albopunctatum	Didemnum cf. albopunctatum SLI-2012 cytochrome	603	KC017444
Aplousobranchia	Didemnidae	Didemnum cineraceum	Didemnum cineraceum voucher ColBIO TL 230 cyto	598	KU221193
Aplousobranchia	Didemnidae	Didemnum cuculliferum	Didemnum cuculliferum voucher UF 724 cytochrome	605	KU221219
Aplousobranchia	Didemnidae	Didemnum fragile	Didemnum fragile voucher UF 839 cytochrome oxid	605	KU221201
Aplousobranchia	Didemnidae	Didemnum fulgens	Didemnum fulgens cytochrome oxidase subunit I ge	604	JX846617
Aplousobranchia	Didemnidae	Didemnum granulatum	Didemnum granulatum strain Sni2 cytochrome oxid	603	JQ013198
Aplousobranchia	Didemnidae	Didemnum granulatum	Didemnum granulatum voucher UF 1408 cytochrome	604	KU221215
Aplousobranchia	Didemnidae	Didemnum granulatum	Didemnum granulatum voucher UF 573 cytochrome	605	KU221196
Aplousobranchia	Didemnidae	Didemnum incanum	Didemnum incanum haplotype 8 cytochrome oxidase	594	JQ692626
Aplousobranchia	Didemnidae	Didemnum ligulum	Didemnum ligulum voucher ColBIO TL796 cytochrome	604	KU221210
Aplousobranchia	Didemnidae	Didemnum ligulum	Didemnum ligulum voucher UF 1387 cytochrome oxid	605	KU221202
Aplousobranchia	Didemnidae	Didemnum lutarium	Didemnum lutarium isolate 17Jul14-1-4 cytochrome	598	KY111418
Aplousobranchia	Didemnidae	Didemnum membranaceum	Didemnum membranaceum voucher UF 526 cytochr	605	KU221216
Aplousobranchia	Didemnidae	Didemnum multispirale	Didemnum multispirale cytochrome c oxidase subun	519	KC017430
Aplousobranchia	Didemnidae	Didemnum mutabile	Didemnum mutabile voucher UF 691 cytochrome ox	604	KU221223
Aplousobranchia	Didemnidae	Didemnum patulum	Didemnum patulum haplotype 2 cytochrome oxidase	590	JQ731737
Aplousobranchia	Didemnidae	Didemnum perlicudum	Didemnum perlicudum voucher DBTIC117 cytochr	610	KX138490
Aplousobranchia	Didemnidae	Didemnum psammatoide	Didemnum psammatoide voucher RMNH.UROCH.746	586	EU742661
Aplousobranchia	Didemnidae	Didemnum psammatoides	Didemnum psammatoides voucher ColBIO TL 27 cyt	598	KU221189
Aplousobranchia	Didemnidae	Didemnum sordidum	Didemnum sordidum voucher UF 612 cytochrome o	605	KU221197
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. 1 SLI-2012 cytochrome c oxidase su	605	KC017439
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. 2 SLI-2012 cytochrome c oxidase su	608	KC017432
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. BAOR-2012 isolate E cytochrome oxi	609	JQ731747
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. IEAPMC1237 cytochrome oxidase su	619	KX641024
Aplousobranchia	Didemnidae	Didemnum sp. ?	Tava2	847	This study, LR136924
Aplousobranchia	Didemnidae	Didemnum sp.	Unpublished Tava-3 sample	843	Our unpublished seq, LR136925
Aplousobranchia	Didemnidae	Didemnum sp. B LMS-2	Didemnum sp. B LMS-2008 voucher RMNH UROCH 7	586	EU419408
Aplousobranchia	Didemnidae	Didemnum sp. BAOR-2012	Didemnum sp. BAOR-2012 haplotype 11 cytochrome	512	JQ731741
Aplousobranchia	Didemnidae	Didemnum sp. BAOR-2012	Didemnum sp. BAOR-2012 haplotype 12 cytochrome	560	JQ731742
Aplousobranchia	Didemnidae	Didemnum sp. BAOR-2012	Didemnum sp. BAOR-2012 isolate B cytochrome oxi	569	JQ731744
Aplousobranchia	Didemnidae	Didemnum sp. DS-4	Didemnum sp. DS-4 cytochrome oxidase subunit I (	586	KF309622
Aplousobranchia	Didemnidae	Didemnum sp. ESC-10	Didemnum sp. ESC-10 cytochrome oxidase subunit	586	KF309573
Aplousobranchia	Didemnidae	Didemnum sp. GA-2014	Didemnum sp. GA-2014 cytochrome c oxidase subu	599	KJ725167
Aplousobranchia	Didemnidae	Didemnum vexillum	Didemnum vexillum isolate GAASC3 cytochrome c o	606	KJ725154
Aplousobranchia	Didemnidae	Didemnum vexillum Clade A	entire mtDNA and COI		KM259616
Aplousobranchia	Didemnidae	Didemnum vexillum Clade B	entire mtDNA and COI		KM259617
Aplousobranchia	Didemnidae	Diplosoma listerianum	entire mtDNA and COI		FN313539
Aplousobranchia	Didemnidae	Diplosoma simile	Diplosoma simile mitochondrial coxI gene for cytoch	631	AB723717
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. 1 AO-2016 voucher ColBIO TL 729 cy	603	KU221184
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. 2 AO-2016 voucher UF 644 cytochrome	604	KU221220
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. 3 AO-2016 voucher UF 680 cytochrome	572	KU221228
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. 3 AO-2016 voucher UF 759 cytochrome	604	KU221226
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. AO-2016 voucher ColBIO TL 817 cyt	596	KU221188
Aplousobranchia	Didemnidae	Diplosoma spongiforme	Diplosoma spongiforme cytochrome c oxidase subun	617	AY600972
Aplousobranchia	Didemnidae	Lissoclinum abdominale	Lissoclinum abdominale voucher ICBT17 02 cytochr	661	KY741542
Aplousobranchia	Didemnidae	Lissoclinum aff. fragile	Lissoclinum aff. fragile SLI-2011 clone 09Jun081-2	605	JF506182
Aplousobranchia	Didemnidae	Lissoclinum fragile	Lissoclinum fragile clone 08Jun08-1-5 cytochrome c	605	JF506180
Aplousobranchia	Didemnidae	Lissoclinum patella	Lissoclinum patella isolate 07-110 cytochrome oxida	895	KJ009374
Aplousobranchia	Didemnidae	Lissoclinum patella L2	entire mtDNA and COI		KJ596321
Aplousobranchia	Didemnidae	Lissoclinum patella L5	entire mtDNA and COI		KJ596322
Aplousobranchia	Didemnidae	Lissoclinum patella L6	entire mtDNA and COI		KJ596323
Aplousobranchia	Didemnidae	Lissoclinum ravarava	Lissoclinum ravarava voucher UF 1398 cytochrome c	568	KU221214
Aplousobranchia	Didemnidae	Lissoclinum vareau	Lissoclinum vareau voucher UF 606 cytochrome oxid	601	KU221205
Aplousobranchia	Didemnidae	Lissoclinum verrilli	Lissoclinum verrilli clone ZPLV cytochrome c oxidase	511	JX099361
Aplousobranchia	Didemnidae	Polysyncraton lacazei	Polysyncraton lacazei cytochrome c oxidase subunit	617	AY600986
Aplousobranchia	Didemnidae	Polysyncraton poro	Polysyncraton poro voucher UF 664 cytochrome oxid	605	KU221217
Aplousobranchia	Didemnidae	Polysyncraton sp.	Polysyncraton sp. AO-2016 voucher ColBIO TL 432	598	KU221187
Aplousobranchia	Didemnidae	Polysyncraton sp.	Polysyncraton sp. AO-2016 voucher ColBIO TL 810	593	KU221186
Aplousobranchia	Didemnidae	Trididemnum cereum	Trididemnum cereum isolate PRE-AA cytochrome ox	586	KF309632
Aplousobranchia	Didemnidae	Trididemnum cyanophorum	Trididemnum cyanophorum clone 30May08-2-1 cyt	605	JF506187
Aplousobranchia	Didemnidae	Trididemnum fetia	Trididemnum fetia voucher UF 629 cytochrome oxid	604	KU221206
Aplousobranchia	Didemnidae	Trididemnum maragogi	Trididemnum maragogi voucher ColBIO TL 234 cyto	603	KU221183
Aplousobranchia	Didemnidae	Trididemnum pigmentatum	Trididemnum pigmentatum voucher UF 696 cytochr	604	KU221224
Aplousobranchia	Didemnidae	Trididemnum solidum	Trididemnum solidum clone 9Jun08-2-1 cytochrome	587	JF506186
Aplousobranchia	Didemnidae	Trididemnum sp.	Trididemnum sp. AO-2016 voucher ColBIO TL 538 d	594	KU221185
Aplousobranchia	Didemnidae	Trididemnum tomarahi	Trididemnum tomarahi voucher UF 1424 cytochrome	605	KU221203
Aplousobranchia	Holozoidae	Distaplia bermudensis	Distaplia bermudensis isolate 1D cytochrome oxidase	652	MG525009
Aplousobranchia	Polycitoridae	Cystodytes dellechiaiei	Cystodytes dellechiaiei haplotype S1 cytochrome oxi	617	AY523042
Aplousobranchia	Polycitoridae	Cystodytes sp.	Cystodytes sp. SLI-2005 cytochrome oxidase subun	617	AY821784
Aplousobranchia	Polycitoridae	Eudistoma amplum	Eudistoma amplum cytochrome c oxidase subunit I	605	KC017442
Aplousobranchia	Polycitoridae	Eudistoma banyulensis	Eudistoma banyulensis cytochrome c oxidase subun	617	AY600973
Aplousobranchia	Polycitoridae	Eudistoma capsulatum	Eudistoma capsulatum isolate 5A-1 cytochrome oxid	659	MG525019
Aplousobranchia	Polycitoridae	Eudistoma elongatum	Eudistoma elongatum mitochondrial partial COI gen	702	HG005369
Aplousobranchia	Polycitoridae	Eudistoma microlarvum	Eudistoma microlarvum voucher DBTIC18 cytochrome	544	KM411614
Aplousobranchia	Polycitoridae	Eudistoma ovatum	Eudistoma ovatum voucher DBTIC20 cytochrome ox	583	KM411610
Aplousobranchia	Polycitoridae	Eudistoma planum	Eudistoma planum cytochrome c oxidase subunit I (	617	AY600977
Aplousobranchia	Polycitoridae	Eudistoma plumbeum	Eudistoma plumbeum cytochrome c oxidase subunit	617	AY600978
Aplousobranchia	Polycitoridae	Eudistoma posidoniarum	Eudistoma posidoniarum cytochrome c oxidase subu	617	AY600974
Aplousobranchia	Polycitoridae	Eudistoma viride	Eudistoma viride cytochrome c oxidase subunit I (C	608	KY043421
Aplousobranchia	Polycitoridae	Nephtheis fascicularis	Nephtheis fascicularis coi gene for cytochrome c oxi	629	AM706489
Aplousobranchia	Polycitoridae	Polycitor adriaticum	Polycitor adriaticum cytochrome c oxidase subunit I	617	AY600982
Aplousobranchia	Polycitoridae	Polycitor giganteus	Polycitor giganteus cytochrome c oxidase subunit I (	621	KC017429
Aplousobranchia	Polycitoridae	Polycitorella coronaria	Polycitorella coronaria mitochondrial partial COI gen	636	AM709504
Aplousobranchia	Polycitoridae/Pseudodistomidae	Eudistoma or Pseudodistoma	Tava8	822	This study, LR136942
Aplousobranchia	Apolidium cf. accarens	Apolidium cf. accarens	Apolidium cf. accarens ESC-6 cytochrome oxidase su		KF309571
Aplousobranchia	Polyclinidae	Apolidium coeruleum	partial mtDNA, entire COI		published in PMID:24572017
Aplousobranchia	Polyclinidae	Apolidium conicum	entire mtDNA and COI		FN313538
Aplousobranchia	Polyclinidae	Apolidium elegans	Apolidium elegans cytochrome c oxidase subunit I (C	617	AY600971
Aplousobranchia	Polyclinidae	Apolidium fuscum	Apolidium fuscum cytochrome c oxidase subunit I (C	617	AY600975
Aplousobranchia	Polyclinidae	Apolidium pseudolobatum	Apolidium pseudolobatum cytochrome c oxidase subu	617	AY600967
Aplousobranchia	Polyclinidae	Apolidium stellatum	Amarocium stellatum cytochrome oxidase subunit I	670	AY116595
Aplousobranchia	Polyclinidae	Apolidium stellatum	Apolidium stellatum isolate 17Jul14-1-3 White cyto	606	KY111411
Aplousobranchia	Polyclinidae	Apolidium tabarquensis	partial mtDNA and COI		HF548555
Aplousobranchia	Polyclinidae	Morchellium argus	Morchellium argus isolate DS-2 cytochrome oxidase	586	KF309620
Aplousobranchia	Polyclinidae	Polyclinium indicum	Polyclinium indicum voucher DBTIC157 cytochrome c	667	KX650777
Aplousobranchia	Polyclinidae	Synoicum castellatum	Synoicum castellatum cytochrome c oxidase subunit	621	KC017441
Aplousobranchia	Polyclinidae	Synoicum castellatum	Synoicum castellatum isolate GAASC10 cytochrome	617	KJ725161
Aplousobranchia	Pseudodistomidae	Pseudodistoma crucigaster	Pseudodistoma crucigaster cytochrome c oxidase su	617	AY600979
Aplousobranchia	Pseudodistomidae	Pseudodistoma crucigaster	Pseudodistoma crucigaster mitochondrial partial coi	578	AJ748711
Aplousobranchia	Pseudodistomidae	Pseudodistoma cymusense	Pseudodistoma cymusense cytochrome c oxidase su	617	AY600970
Aplousobranchia	Pyncoclavellidae	Pyncoclavella atlantica	Pyncoclavella atlantica mitochondrial coi gene for cy	599	AM403685
Aplousobranchia	Pyncoclavellidae	Pyncoclavella aurilucens	Pyncoclavella aurilucens mitochondrial coi gene for c	610	AM403692
Aplousobranchia	Pyncoclavellidae	Pyncoclavella brava	Pyncoclavella brava mitochondrial coi gene for cyto	610	AM403688
Aplousobranchia	Pyncoclavellidae	Pyncoclavella communis	Pyncoclavella communis mitochondrial partial coi ge	702	AM746353
Aplousobranchia	Pyncoclavellidae	Pyncoclavella diminuta	Pyncoclavella diminuta cytochrome c oxidase subun	621	KC017435
Aplousobranchia	Pyncoclavellidae	Pyncoclavella martae	Pyncoclavella martae coi gene for cytochrome c oxid	628	AM706481
Aplousobranchia	Pyncoclavellidae	Pyncoclavella nana	Pyncoclavella nana mitochondrial coi gene for cytoch	610	AM403699
Aplousobranchia	Pyncoclavellidae	Pyncoclavella producta (erroneous assignment)	Archidistoma aggregatum cytochrome c oxidase sub	537	AY600966
Aplousobranchia	Pyncoclavellidae	Pyncoclavella sp. RPP-2006	Pyncoclavella sp. RPP-2006 mitochondrial coi gene f	610	AM403704
Aplousobranchia	Pyncoclavellidae	Pyncoclavella sp. RPP-2006	Pyncoclavella sp. RPP-2006 mitochondrial coi gene f	587	AM403706
Aplousobranchia	Pyncoclavellidae	Pyncoclavella sp. SLI-2004	Pyncoclavella sp. SLI-2004 cytochrome c oxidase su	617	AY600988
Aplousobranchia	Pyncoclavellidae	Pyncoclavella sp. SLI-2012	Pyncoclavella sp. SLI-2012 cytochrome c oxidase su	620	KC017436
Aplousobranchia	Pyncoclavellidae	Pyncoclavella tabella	Pyncoclavella tabella coi gene for cytochrome c oxid	594	AM706488
Phlebobranchia	Ascididae	Ascidella aspersa	entire mtDNA and COI		HF548561
Phlebobranchia	Ascididae	Phallusia fumigata	entire mtDNA and COI		AM292602
Phlebobranchia	Ascididae	Phallusia mammillata	entire mtDNA and COI		AM292320
Phlebobranchia	Cionidae	Ciona intestinalis (formerly Ciona intestinalis)	entire mtDNA and COI		AM292218
Phlebobranchia	Cionidae	Ciona robusta (formerly Ciona intestinalis type)	entire mtDNA and COI		AJ517314
Phlebobranchia	Cionidae	Ciona savignyi	entire mtDNA and COI		AB079784
Stolidobranchia	Pyluridae	Halocynthia papillosa	entire mtDNA and COI		FM177863
Stolidobranchia	Pyluridae	Halocynthia roretzi	entire mtDNA and COI		AB024528
Stolidobranchia	Pyluridae	Halocynthia spinosa	entire mtDNA and COI		HF548558
Stolidobranchia	Pyluridae	Herdmania morius	entire mtDNA and COI		FN296153
Stolidobranchia	Pyluridae	Microcosmus sulcatus	entire mtDNA and COI		AM292321
Stolidobranchia	Pyluridae	Pyura ganglion	entire mtDNA and COI		HF548557
Stolidobranchia	Styelidae	Botrylloides leachi	entire mtDNA and COI		HG931921
Stolidobranchia	Styelidae	Botrylloides nigrum	entire mtDNA and COI		HF548559
Stolidobranchia	Styelidae	Botrylloides pizoni	entire mtDNA and COI		HF548554
Stolidobranchia	Styelidae	Botrylloides violaceus	entire mtDNA and COI		HF548552
Stolidobranchia	Styelidae	Botryllus schlosseri	entire mtDNA and COI		FM177702
Stolidobranchia	Styelidae	Polycarpa mytiligera	entire mtDNA and COI		HF548556
Stolidobranchia	Styelidae	Styela clava	entire mtDNA and COI		HG931920
Stolidobranchia	Styelidae	Styela plicata	entire mtDNA and COI		AM292601

Table S64. List of the COI sequences included in our phylogenetic reconstructions. For each sequence is reported the Accession Number (AC), the description line, the species name, the order/family and the sequence length.

Order	Family	Species	Description Line	bp	AC
Aplousobranchia	Clavelinidae	Clavelina meridionalis	Clavelina meridionalis partial 18S r	1718	FM244840
Aplousobranchia	Didemnidae	Didemnum albidum	Didemnum albidum isolate GBAldid	1782	EU337059
Aplousobranchia	Didemnidae	Didemnum albidum	Didemnum albidum isolate GBAldid	1798	EU337058
Aplousobranchia	Didemnidae	Didemnum cf. ligulum	Didemnum cf. ligulum AO-2016 vou	1041	KU170205
Aplousobranchia	Didemnidae	Didemnum molle	Didemnum molle gene for 18S rRNA	2262	AB211071
Aplousobranchia	Didemnidae	Didemnum sp?	Tava2	559	This study, LR136919
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. A GBsp.A/5 18S rib	1970	EU337060
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. B NZsp.B17 18S rib	2002	EU337061
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. DidSA/57 gene for 1	2248	AB211072
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. DidSB gene for 18S	2010	AB211073
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. Mahanga Bay-EM-20	938	AJ579862
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. USA-EM-2003 18S r	922	AJ579863
Aplousobranchia	Didemnidae	Didemnum sp.	Didemnum sp. Whangamata-EM-20	2197	AJ579861
Aplousobranchia	Didemnidae	Didemnum vexillum	Didemnum vexillum partial 18S rRNA	643	HG005366
Aplousobranchia	Didemnidae	Didemnum vexillum Clade A	Didemnum vexillum haplotype Clad	861	JF738070
Aplousobranchia	Didemnidae	Didemnum vexillum Clade B	Didemnum vexillum haplotype Clad	861	JF738071
Aplousobranchia	Didemnidae	Diplosoma mitsukurii	Diplosoma mitsukurii gene for 18S	2084	AB211096
Aplousobranchia	Didemnidae	Diplosoma ooru	Diplosoma ooru gene for 18S rRNA	2081	AB211097
Aplousobranchia	Didemnidae	Diplosoma simile	Diplosoma simile 18S ribosomal RN	2032	KM458192
Aplousobranchia	Didemnidae	Diplosoma simile	Diplosoma simile gene for 18S rRNA	2075	AB211101
Aplousobranchia	Didemnidae	Diplosoma simileguwa	Diplosoma simileguwa gene for 18S	2075	AB211107
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. AO-2016 voucher Co	884	KU170204
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. DipSS/111 gene for	2074	AB211121
Aplousobranchia	Didemnidae	Diplosoma sp.	Diplosoma sp. DipVS1/107 gene for	2074	AB211116
Aplousobranchia	Didemnidae	Diplosoma virens	Diplosoma virens gene for 18S rRNA	2073	AB211109
Aplousobranchia	Didemnidae	Leptoclinides madara	Leptoclinides madara gene for 18S	2042	AB211070
Aplousobranchia	Didemnidae	Lissoclinum badium	Lissoclinum badium gene for 18S r	1962	AB211078
Aplousobranchia	Didemnidae	Lissoclinum bistratum	Lissoclinum bistratum gene for 18S	1916	AB211079
Aplousobranchia	Didemnidae	Lissoclinum patella	Lissoclinum patella gene for 18S r	1916	AB211085
Aplousobranchia	Didemnidae	Lissoclinum punctatum	Lissoclinum punctatum gene for 18	1916	AB211086
Aplousobranchia	Didemnidae	Lissoclinum sp.	Lissoclinum sp. LisS/94 gene for 18	1916	AB211095
Aplousobranchia	Didemnidae	Lissoclinum timorense	Lissoclinum timorense gene for 18S	1916	AB211092
Aplousobranchia	Didemnidae	Polysyncrator	Polysyncrator sp. AO-2016 voucher	919	KU170203
Aplousobranchia	Didemnidae	Trididemnum cyanophorum	Trididemnum cyanophorum strain M	1776	KJ632950
Aplousobranchia	Didemnidae	Trididemnum fetia	Trididemnum fetia voucher UF_629	1031	KU174392
Aplousobranchia	Didemnidae	Trididemnum maragogi	Trididemnum maragogi voucher Col	862	KU170201
Aplousobranchia	Didemnidae	Trididemnum paracyclops	Trididemnum paracyclops gene for	2108	AB211074
Aplousobranchia	Didemnidae	Trididemnum tomarahi	Trididemnum tomarahi voucher UF	1008	KU174391
Aplousobranchia	Holozoidae	Distaplia dubia	Distaplia dubia gene for 18S rRNA	1782	AB211068
Aplousobranchia	Holozoidae	Distaplia sp.	Distaplia sp. AO-2016 18S ribosom	908	KU174394
Aplousobranchia	Polycitoridae	Cystodytes sp.	Cystodytes sp. TG-2008 partial 18S	2479	FM244842
Aplousobranchia	Polycitoridae	Eudistoma elongatum	Eudistoma elongatum partial 18S r	589	HG005367
Aplousobranchia	Polycitoridae	Eudistoma gilboviride	Eudistoma gilboviride gene for 18S	1833	AB211069
Aplousobranchia	Polycitoridae/Pseudodistomidae	Eudistoma or Pseudodistoma	Tava8	566	This study, LR136920
Aplousobranchia	Polyclinidae	Aplidium conicum	Aplidium conicum strain MH-JMC5	1794	KJ632949
Aplousobranchia	Polyclinidae	Aplidium pliciferum	Aplidium pliciferum gene for 18S r	1902	AB211067
Aplousobranchia	Pycnoclavellidae	Pycnoclavella aff. detorta	Pycnoclavella aff. detorta TG-2008	1725	FM244841
Aplousobranchia	Pycnoclavellidae	Pycnoclavella diminuta	Pycnoclavella diminuta strain MH-JN	1731	KJ632948
Phlebobranchia	Cionidae	Ciona intestinalis	Ciona intestinalis gene for 18S rRNA	1756	AB013017
Phlebobranchia	Cionidae	Ciona savignyi	Uncultured eukaryote gene for smal	1765	AB191429

Table S65. List of the 18S rDNA sequences included in our phylogenetic reconstructions. For each sequence is reported the Accession Number (AC), the description line, the species name, the order/family and the sequence length.

Order	Family	Genus	Genus Abbr.
Aplousobranchia	Clavelinidae	Clavelina	Cl
	Didemnidae	Didemnum	Dd
		Diplosoma	Dp
		Lissoclinum	Ls
		Polysyncrator	Pls
		Trididemnum	Tr
	Holozoidae	Distaplia	Dt
	Polycitoridae	Cystodytes	Cs
		Eudistoma	Ed
		Nephtheis	Np
		Polycitor	Pl
		Polycitorella	Pll
	Polyclinidae	Aplidium	Ap
		Morchellium	Mr

		Polyclinum	Pln
		Synoicum	Sn
	Pycnoclavellidae	Pycnoclavella	Py
	Pseudodistomidae	Pseudodistoma	Ps
	Diazonidae	Rhopalaea	Rh
Phlebobranchia	Cionidae	Ciona	Cn
	Asciidiidae	Ascidia	As
		Phallusia	Ph
Stolidobranchia	Pyuridae	Halocynthia	Hl
		Herdmania	Hr
		Microcosmus	Mc
		Pyura	Py
	Styelidae	Botrylloides	Btl
		Botryllus	Bt
		Polycarpa	Pc
		Styela	St

Table S66. Abbreviation of the genera analysed.

REFERENCES

- (1) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminformatics* **2012**, *4*, 17.
- (2) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (3) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. *J. Cheminformatics* **2011**, *3*, 33.
- (4) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396. h
- (5) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (6) Cammi, R.; Tomasi, J. *Int. J. Quantum Chem.* **1995**, *56*, 465–474.
- (7) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.
- (8) Rappoport, D.; Furche, F. *J. Chem. Phys.* **2010**, *133*, 134105.
- (9) Neese, F.. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (10) ACD/Labs.com :: Your Partner in Chemistry Software for Analytical and Chemical Knowledge Management, Chemical Nomenclature, and In-Silico PhysChem and ADME-Tox <http://www.acdlabs.com/home/> (accessed Mar 9, 2017).
- (11) Hirose, M.; Hirose, E. *Zoolog. Sci.* **2009**, *26*, 564–568.
- (12) Brunetti, R.; Manni, L.; Mastrototaro, F.; Gissi, C.; Gasparini, F. *Zootaxa* **2017**, *4353*, 29.
- (13) Wada, H.; Makabe, K. W.; Nakauchi, M.; Satoh, N. *Biol. Bull.* **1992**, *183*, 448–455.
- (14) Tsagkogeorga, G.; Turon, X.; Hopcroft, R. R.; Tilak, M.-K.; Feldstein, T.; Shenkar, N.; Loya, Y.; Huchon, D.; Douzery, E. J.; Delsuc, F. *BMC Evol. Biol.* **2009**, *9*, 187.
- (15) Yokobori, N.; López, B.; Monteserin, J.; Paul, R.; Von Groll, A.; Martin, A.; Marquina-Castillo, B.; Palomino, J. C.; Hernández-Pando, R.; Sasiain, M. del C.; et al. *Int. J. Med. Microbiol.* **2018**, *308*, 349–357.
- (16) Takara Bio—Home <https://www.takarabio.com/> (accessed Dec 15, 2018).
- (17) Thermo Fisher Scientific - UK <https://www.thermofisher.com/uk/en/home.html> (accessed Dec 15, 2018).
- (18) ZYMO RESEARCH | DNA and RNA Purification Products and Technology Leader <https://www.zymoresearch.eu/> (accessed Dec 15, 2018).

- (19) Eurofins Genomics Munich <https://www.eurofins.com/contact-us/worldwide-interactive-map/germany/eurofins-genomics-munich/> (accessed Dec 15, 2018).
- (20) Microsynth AG | swissbiotech.org <https://www.swissbiotech.org/member/microsynth-ag> (accessed Dec 15, 2018).
- (21) Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. *Bioinformatics* **2012**, *28*, 1647–1649.
- (22) Home - Nucleotide - NCBI <https://www.ncbi.nlm.nih.gov/nucleotide/> (accessed Dec 15, 2018).
- (23) Nucleotide BLAST: Search nucleotide databases using a nucleotide query https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch (accessed Dec 15, 2018).
- (24) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403–410.
- (25) Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. *Nucleic Acids Res.* **2002**, *30*, 3059–3066.
- (26) Tsagkogeorga, G.; Turon, X.; Hopcroft, R. R.; Tilak, M.-K.; Feldstein, T.; Shenkar, N.; Loya, Y.; Huchon, D.; Douzery, E. J. P.; Delsuc, F. Tsagkogeorga-BMCEvolBiol2010-Tunicata_18S_88taxa. **2011**. <https://doi.org/10.5061/dryad.8333/1>.
- (27) Katoh, K.; Rozewicki, J.; Yamada, K. D. MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization. *Brief. Bioinform.* **2017**. <https://doi.org/10.1093/bib/bbx108>.
- (28) ATGC: PhyML with automatic model selection by SMS <http://www.atgc-montpellier.fr/phyml-sms/> (accessed Dec 15, 2018).
- (29) Guindon, S.; Gascuel, O. *Syst. Biol.* **2003**, *52*, 696–704.
- (30) Brunetti, R.; Manni, L.; Mastrototaro, F.; Gissi, C.; Gasparini, F. *Zootaxa* **2017**, *4353*, 29.
- (31) Hebert, P. D. N.; Cywinska, A.; Ball, S. L.; deWaard, J. R. *Proc. R. Soc. B Biol. Sci.* **2003**, *270*, 313–321.
- (32) Tianero, M. D.; Pierce, E.; Raghuraman, S.; Sardar, D.; McIntosh, J. A.; Heemstra, J. R.; Schonrock, Z.; Covington, B. C.; Maschek, J. A.; Cox, J. E.; et al. *Proc. Natl. Acad. Sci.* **2016**, *113*, 1772–1777.
- (33) da Silva Oliveira, F. A.; Michonneau, F.; da Cruz Lotufo, T. M. *Zool. J. Linn. Soc.* **2017**, *180*, 603–612.
- (34) Moreno, T. R.; Rocha, R. M. *Rev. Bras. Zool.* **2008**, *25*, 269–298.
- (35) Turon, X.; López-Legentil, S. *Mol. Phylogenet. Evol.* **2004**, *33*, 309–320.